ACIC REFERENCE PUBLICATION NO. 16

ERROR ANALYSIS BY THE COVARIANCE METHOD

JANUARY 1963

Aeronautical Chart and Information Center
United States Air Force
St. Louis 18, Missouri

ERROR ANALYSIS BY THE COVARIANCE METHOD

ACIC Reference Publication No. 16

January 1963

Prepared by
Melvin E. Shultz
Donald A. Richardson
Geophysical Studies Section
Geo-Science Branch
Chart Research Division

ABSTRACT. For the most part, error analysis has been treated as though only independent errors are involved. For convenience and simplicity, this has been a valid assumption in error analysis at ACIC. However, with the necessity for analyzing and reducing massive data an awareness of the existence of dependent errors and their significance in error analysis becomes increasingly necessary.

The analysis of dependent errors makes use of the concept of distribution moments and the moment matrix (covariance matrix). This paper presents an analysis of the normal bivariate and trivariate error distributions along with their relationships to the moment matrix, and the application of this concept to least squares and adjustments. It is shown that a suitable transformation of the covariance matrix yields independent errors which may be substituted for the dependent errors in further error analysis. A brief introduction to matrix properties is also included for background information.

AERONAUTICAL CHART AND INFORMATION CENTER
Second and Arsenal Streets
St. Louis 18, Missouri

This report is issued to disseminate timely, useful technical procedure to activities engaged in geodesy and related subjects. Nothing in this report is to be construed as necessarily coinciding with United States Air Force Doctrine.

This report was prepared under the direction of the Aeronautical Chart and Information Center, USAF. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CONTENTS

LIST OF SYMBOLS

x	=	a random variable
x_i	=	the i^{th} value of the random variable x
$\bar{x}$	=	mean value of the variable x
$E\{x\}$	=	expected value of the variable x
$\sum_{i=1}^{n} x_i$	=	summation of all values of x from x_1 to x_n, that is $x_1 + x_2 + x_3 + \ldots x_n$
n	=	number of measurements, generally refers to the last measurement (the n^{th} measurement)
σ	=	standard error
σ_x, σ_y	=	standard error of x and y respectively
σ^2	=	variance
σ_x^2, σ_y^2	=	variance of x and y respectively
σ_{xy}	=	covariance of x with y
σ_{yx}	=	covariance of y with x
ρ	=	correlation coefficient
μ_x, μ_y	=	first moments of x and y respectively (theoretical mean values)
μ_{xx}, μ_{yy}	=	second moment of x and y respectively (theoretical variance)
$\mu_{\bar{x}\bar{x}}$, $\mu_{\bar{y}\bar{y}}$	=	second moment about the mean value (central moments)
μ_{xy}, μ_{yx}	=	mixed product moment (theoretical covariance)

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \text{the A matrix}$$

a_{ij} = the element in the i^{th} row and j^{th} column of matrix A

$$I = \text{unit matrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

0 = null matrix, all elements equal to zero

$|A|$ = determinant of a square matrix A

A^T = transposed matrix of A

A^T = $[a_{ji}]$

A^{-1} = inverse matrix of A

a^{ij} = cofactor of a_{ij} element

$\frac{dA(x)}{dx}$ = derivative of matrix A with respect to x

$\frac{\partial A(x)}{\partial x_1}$ = partial derivative of matrix A with respect to x

$E\left\{e^{\theta x}\right\}$ = moment generating function (univariate case)

$E\left\{\exp\,[\theta x + \theta y]\right\}$ = moment generating function (bivariate case)

$$M = \text{moment matrix} = \begin{bmatrix} \mu_{xx} & \mu_{xy} & \cdots \\ \mu_{yx} & \mu_{yy} & \cdots \\ \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \cdots \\ \sigma_{yx} & \sigma_y^2 & \cdots \\ \vdots & \vdots & \end{bmatrix}$$

$f = f(x,y)$ = normal distribution function (bivariate case)

$f = f(x,y,z)$ = normal distribution function (trivariate case)

Q = general quadratic equation

Q_2 = quadratic expression of the normal bivariate distribution

Q_3 = quadratic expression of the normal trivariate distribution

K = constant of general normal distribution

K_2 = constant of bivariate normal distribution

K_3 = constant of trivariate normal distribution

D = distance between center of ellipse and a point on its perimeter

F = function involved in Lagrangian multiplier

λ = latent root

g = side condition of Lagrangian multiplier

ξ = true value of measured quantity

E = true value matrix

ϵ = true error

ϵ' = true error matrix

v = residuals

V = residual matrix

k_i = constants from set of assumed observations

s_i = constants from theory of the observation

Z_i = value observed

P = weight matrix

B = $A \cdot P \cdot K$

C = $A \cdot P \cdot K$ (introduced for B matrix)

X' = matrix set of x_1', x_2', ..., x_n' variables

V' = residual matrix of x_1', x_2', ..., x_n' variables

$[pvv]$ = sum of squares of the weighted residuals

1. INTRODUCTION

1.1. Purpose. The occurrence of data involving several variables is quite common to the reduction of geodetic information to a useful form. Reduction of such information requires certain generalizations and, for simplicity, it is often assumed that the variables are completely independent of each other. In many instances, this is not a valid assumption and some type of reduction which considers the interrelationship between variables is necessary.

The following sections describe a method of reduction which will reflect the correlation between the variables. Extensive use is made of statistical, geometrical, and matrix properties.

1.2. Assumptions and Generalizations. The mean value ($\bar{x}$) of a random variable is computed by:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

where:

$\bar{x}$ = mean value

x = a random variable

$\sum_{i=1}^{n}$ = sign indicating the sum of all values $(x_1, x_2, \ldots, x_n)$

n = number of measurements.

As n grows larger the value of $\bar{x}$ becomes more reliable. In theoretical statistics, n is considered to approach infinity and therefore, incorporates a universe of readings. When n approaches ∞, the computed value approaches the theoretical value. The theoretical value for the mean is generally considered the expected value and symbolized by $E\{x\}$.

Another statistical parameter, the standard error (σ), denotes the dispersion of the measured values around the mean and is computed by:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

The square of the standard error (σ^2) is known as the variance.

Since $\bar{x}$ and n also enter the computation of the standard error, it is again assumed that the true standard error is obtained with n approaching ∞. While the distribution of x is discrete no devastating error results when the discrete function is considered continuous with n ranging from $-\infty$ to $+\infty$. It is also assumed that the variables are normally distributed with systematic errors eliminated.

1.3. Independent and Dependent Errors. In error terminology two types of errors are considered — dependent and independent. The treatment of independent errors is comparatively simple because, as the name implies,they are independent of each other. For example, the magnitude of ϵ_1^I of one variable is completely independent of the magnitude of the error ϵ_2^I of another variable. Dependent errors (ϵ_1^D, ϵ_2^D), however, are more involved as the magnitude and even the occurrence of ϵ_1^D depends on the magnitude or occurrence of ϵ_2^D.

Correlation deals with the interrelationship of dependent errors. The degree of correlation is expressed by a numerical quantity called the correlation coefficient (ρ). Although independent[1] errors are uncorrelated the converse is not necessarily true. It does not follow that two errors are independent solely because their correlation coefficient is zero. Consider the distribution which has a frequency function expressed as $x^2 + y^2 = R^2$. Since this is a circle possessing symmetry, the mean values and ρ would be considered zero and uncorrelated. However, to be independent it is necessary that $f(x,y)$ be of the form $f(x)\ f(y)$.

2. MATRIX PROPERTIES

The matrix method lends itself to problems involving numerous variables and mathematical manipulations. It provides a system of compact expressions for series of equations and permits performance of mathematical processes with a minimum of work. Some of the basic properties of the matrix method are presented at this point to insure uniform understanding.

2.1. Definitions and Notations. An aggregate of numbers arranged in a rectangular array is called a rectangular matrix, or simply, a

1. The theory of independent errors, their probabilities and propagational methods for univariate, bivariate, and trivariate cases is developed in TR-96 "Principles of Error Theory and Cartographic Applications."

matrix. The array has m rows and n columns and is denoted by:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

the first subscript designates the row, the second the column in which the element is located. If the dimensions, m x n (m by n), of A are borne in mind, it suffices to use the brief notation:

$$A = [a_{ij}] \qquad (i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n).$$

Two matrices are equal if their corresponding elements are equal. Matrices composed of a single row are called row matrices and those composed of a single column column matrices. In a square matrix of the n^{th} order, the number of rows equals the number of columns.

Among square matrices, an important role is played by diagonal matrices, i.e., a matrix in which only the elements of the principal (leading) diagonal are different from zero:

$$\begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix}$$

If all the numbers α_i of such a matrix are equal to each other, the matrix is said to be a scalar:

$$\begin{bmatrix} \alpha & 0 & \dots & 0 \\ 0 & \alpha & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \alpha \end{bmatrix}$$

and, if $\alpha = 1$, the matrix is said to be the unit matrix I:

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

A matrix with elements equal to zero is a null or zero matrix and is designated by the symbol 0.

The determinant of a matrix is associated with any square matrix. The determinant of the square matrix P is designated $|P|$ and its elements are the elements of matrix P without disarrangement. If $|P| \neq 0$, then P is called a non-singular matrix.

When the rows and columns in matrix A are interchanged:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [a_{ij}],$$

the <u>transposed matrix</u> or transpose:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ & & \cdots & \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} = [a_{ji}]$$

is obtained.

2.2. <u>Elementary Operations on Matrices</u>. To effectively use matrices, the methods of addition and multiplication must be defined. When the elements of a matrix are obtained by multiplying all the elements of the matrix A by a number α, the result is called the product of α and the matrix A.

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ & & \cdots & \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}$$

For example let:

$$\alpha = 3, \quad \text{and} \quad A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

then:

$$\alpha A = \begin{bmatrix} 3\cdot 3 & 3\cdot 4 & 3\cdot 1 \\ 3\cdot 2 & 3\cdot 6 & 3\cdot 2 \\ 3\cdot 1 & 3\cdot 0 & 3\cdot 1 \end{bmatrix} = \begin{bmatrix} 9 & 12 & 3 \\ 6 & 18 & 6 \\ 3 & 0 & 3 \end{bmatrix}$$

When the elements of a matrix C are the sums of the corresponding elements of A and B, the result is the sum of A and B and is defined only when A and B have the same dimensions.

$$C = A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ & & \cdots & \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

For example, let:

$$A = \begin{bmatrix} 2 & 0 & 0 & 4 \\ 6 & 3 & 2 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 5 & 9 & 6 \\ 3 & -4 & 7 & 8 \end{bmatrix}$$

Then:

$$A + B = C = \begin{bmatrix} 2+1 & 0+5 & 0+9 & 4+6 \\ 6+3 & 3+(-4) & 2+7 & 1+8 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 9 & 10 \\ 9 & -1 & 9 & 9 \end{bmatrix}$$

2.2.1. Basic Properties of Matrix Addition and Multiplication.

$$A + (B + C) = (A + B) + C$$

$$A + B = B + A$$

$$A + 0 = A$$

$$(\alpha + \beta) A = \alpha A + \beta A$$

$$\alpha (A + B) = \alpha A + \alpha B$$

where α and β are numbers.

2.2.2. Multiplication of Matrices. Multiplication of the matrices A and B is defined if, and only if, the number of columns of matrix A equals the number of rows of matrix B. The elements of the product, C = AB, are defined in the following manner: the element in the i^{th} row and j^{th} column of C is equal to the sum of the products of the elements of the i^{th} row of matrix A by the corresponding elements of the j^{th} column of matrix B. Thus:

$$C = AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ & & \cdots & \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ & & \cdots & \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

where:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

$$(i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, p)$$

Note that the product of two rectangular matrices is again a rectangular matrix, where the number of rows is equal to the number of rows in the first matrix, and the number of columns equals the number of columns in the second. For example, let:

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 2 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 5 & 2 \\ 6 & 3 \\ 2 & 1 \end{bmatrix}$$

Then:

$$AB = C = \begin{bmatrix} 2\cdot5 + 3\cdot6 + 4\cdot2, & 2\cdot2 + 3\cdot3 + 4\cdot1 \\ 5\cdot5 + 2\cdot6 + 1\cdot2, & 5\cdot2 + 2\cdot3 + 1\cdot1 \end{bmatrix}$$

$$= \begin{bmatrix} 36 & 17 \\ 39 & 17 \end{bmatrix}$$

Another property of matrix multiplication is:

$$(A\cdot B)^T = B^T\cdot A^T$$

or, the transpose of the product of two matrices is equal to the product of the transpose of each matrix in reverse order. (This property may be generalized to products involving more than two matrices.) For example, let:

$$A = \begin{bmatrix} 3 & 4 \\ 1 & 5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 2 \\ 4 & 3 \end{bmatrix},$$

then:

$$AB = \begin{bmatrix} 16 & 18 \\ 20 & 17 \end{bmatrix} \quad \text{and} \quad (AB)^T = \begin{bmatrix} 16 & 20 \\ 18 & 17 \end{bmatrix}$$

Then:

$$(AB)^T \equiv B^T \cdot A^T \equiv \begin{bmatrix} 0 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 4 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 16 & 20 \\ 18 & 17 \end{bmatrix}$$

2.3. Inversion of a Matrix. The inverse a^{-1}, or the reciprocal of a real number, is defined if $a \neq 0$. There is an analogous operation for square matrices. If A is a non-singular square matrix, then a unique matrix written (A^{-1}) exists with the important property:

$$AA^{-1} = I$$

or:

$$A^{-1}A = I$$

where I is the unit matrix. If it exists, the matrix A^{-1} is called the inverse matrix of A and expressed by:

$$A^{-1} = [B_{ij}]$$

where:

$$[b_{ij}] = \frac{[a^{ij}]^T}{|A|} \text{ or equivalently } b_{ij} = \frac{a^{ji}}{|A|}$$

That is, the b_{ij} element (the element in the i^{th} row and j^{th} column of the inverse matrix) is found by 1^{st} transposing the matrix A; then determining the cofactor (a^{ji}) of each element in the transpose; and dividing each cofactor by the determinant of matrix A. The cofactor is the <u>signed</u> minor <u>determinant</u> obtained by deleting all the elements in the same row and column as the element whose cofactor is desired; if the sum of i + j (row number + column number) is even, the minor determinant is multiplied by a positive one (+1); if i + j is odd the minor determinant is multiplied by negative one (-1). For example, let:

$$A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} = [a_{ij}]; \quad A^T = \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix} = [a_{ji}]$$

$$|A| = 3 - 8 = -5$$

$$a^{11} = (+1)\,(1) = 1$$

$$a^{12} = (-1)\,(4) = -4$$

$$a^{21} = (-1)\,(2) = -2$$

$$a^{22} = (+1)\,(3) = 3$$

where $a^{12} = a^{ji}$, the cofactor of the element in the 1^{st} row and 2^{nd} column of A^T. Since $i + j = 1 + 2 = 3$ multiply minor determinant by (-1).

Hence:

$$b_{11} = \frac{a^{11}}{|A|} = \frac{1}{-5} = -\frac{1}{5}$$

$$b_{12} = \frac{a^{21}}{|A|} = \frac{-4}{-5} = \frac{4}{5}$$

$$b_{21} = \frac{a^{12}}{|A|} = \frac{-2}{-5} = \frac{2}{5}$$

$$b_{22} = \frac{a^{22}}{|A|} = \frac{3}{-5} = -\frac{3}{5}$$

and:

$$A^{-1} = B_{ij} = \begin{bmatrix} -\frac{1}{5} & \frac{4}{5} \\ \frac{2}{5} & -\frac{3}{5} \end{bmatrix}$$

The result can be checked by forming the product:

$$A \cdot A^{-1} = \begin{bmatrix} 3(-\frac{1}{5}) + 4(\frac{2}{5}) & 3(\frac{4}{5}) + 4(-\frac{3}{5}) \\ 2(-\frac{1}{5}) + 1(\frac{2}{5}) & 2(\frac{4}{5}) + 1(-\frac{3}{5}) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Since $A \cdot A^{-1} = I$, the inverse A^{-1} is correct.

The above scheme is impractical for matrices of order greater than two. Consequently, other schemes for the computation of inverses have been devised. For a 3rd order square matrix, the procedure is as follows:

let:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

then:

$$A^{-1} = \begin{bmatrix} l & m & n \\ r & s & t \\ u & v & w \end{bmatrix}$$

where:

$$l = \frac{a_{22}a_{33} - a_{23}a_{32}}{|A|}$$

$$m = \frac{a_{13}a_{32} - a_{12}a_{33}}{|A|}$$

$$n = \frac{a_{12}a_{23} - a_{13}a_{22}}{|A|}$$

$$r = \frac{a_{23}a_{31} - a_{21}a_{33}}{|A|}$$

$$s = \frac{a_{11}a_{33} - a_{13}a_{31}}{|A|}$$

$$t = \frac{a_{13}a_{21} - a_{11}a_{23}}{|A|}$$

$$u = \frac{a_{21}a_{32} - a_{22}a_{31}}{|A|}$$

$$v = \frac{a_{12}a_{31} - a_{11}a_{32}}{|A|}$$

$$w = \frac{a_{11}a_{22} - a_{12}a_{21}}{|A|}$$

2.4. Differentiation of Matrices. Let $A(x)$ be a matrix depending on a numerical variable x so that the elements of $A(x)$ are numerical functions of x. That is:

$$A(x) = \begin{bmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x) \\ & & \cdots & \\ a_{m1}(x) & a_{m2}(x) & \cdots & a_{mn}(x) \end{bmatrix}$$

The derivative of $A(x)$ is:

$$\frac{dA(x)}{dx} = \begin{bmatrix} \frac{da_{11}(x)}{dx} & \frac{da_{12}(x)}{dx} & \cdots & \frac{da_{1n}(x)}{dx} \\ \frac{da_{21}(x)}{dx} & \frac{da_{22}(x)}{dx} & \cdots & \frac{da_{2n}(x)}{dx} \\ & & \cdots & \\ \frac{da_{m1}(x)}{dx} & \frac{da_{m2}(x)}{dx} & \cdots & \frac{da_{mn}(x)}{dx} \end{bmatrix}$$

For example let:

$$A(x) = \begin{bmatrix} 3x^2 & 2x^3 & x^4 \\ x & x^2 & 5x^4 \end{bmatrix},$$

then:

$$\frac{dA(x)}{dx} = \begin{bmatrix} 6x & 6x^2 & 4x^3 \\ 1 & 2x & 20x^3 \end{bmatrix}$$

Partial derivatives are obtained similarly. Suppose that:

$$A \cdot X = \begin{bmatrix} 3 & 2 & 5 \\ 2 & 4 & 1 \\ 4 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 + 2x_2 + 5x_3 \\ 2x_1 + 4x_2 + x_3 \\ 4x_1 + 2x_2 + 2x_3 \end{bmatrix}$$

then:

$$\frac{\partial(A \cdot X)}{\partial x_1} = \begin{bmatrix} 3+0+0 \\ 2+0+0 \\ 4+0+0 \end{bmatrix} ; \qquad \frac{\partial(A \cdot X)}{\partial x_2} = \begin{bmatrix} 0+2+0 \\ 0+4+0 \\ 0+2+0 \end{bmatrix}$$

and

$$\frac{\partial(A \cdot X)}{\partial x_3} = \begin{bmatrix} 0+0+5 \\ 0+0+1 \\ 0+0+2 \end{bmatrix}$$

It can be seen that partial differentiation of the rows of $A \cdot X$ with respect to each x_i yields the i^{th} column of matrix A.

2.5. Quadratic Forms. A polynomial:

$$f = f(x,y) = 5x^2 + 6xy + 8y^2,$$

with real coefficients and every term of degree two in x and y, is a quadratic form in x and y. Quadratic forms play a prominent role in analytical geometry and statistics. To relate them to matrices, the polynomial is written:

$$\begin{aligned} f &= 5x^2 + 3xy + 3xy + 8y^2 \\ &= x(5x + 3y) + y(3x + 8y) \\ &= [x,y] \begin{bmatrix} 5x + 3y \\ 3x + 8y \end{bmatrix} = [x,y] \begin{bmatrix} 5 & 3 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

Thus the matrix

$$A = \begin{bmatrix} 5 & 3 \\ 3 & 8 \end{bmatrix}$$

may be used in expressing the quadratic form f:

$$f = X^T A X$$

where:

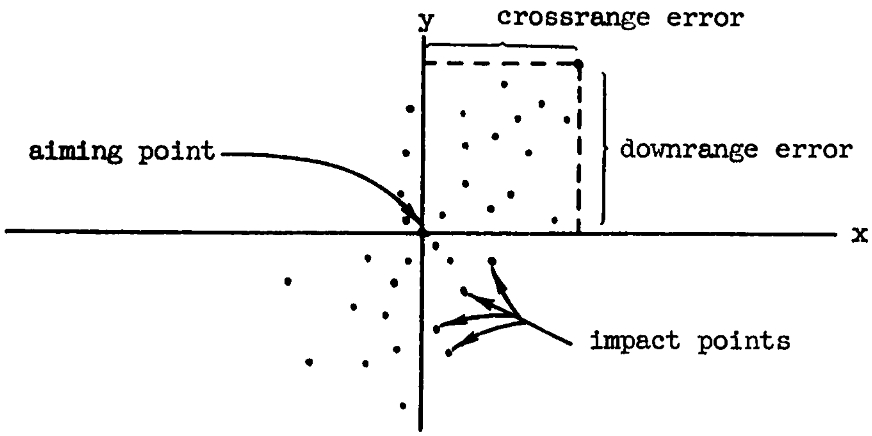
$$X = \text{column matrix} \begin{bmatrix} x \\ y \end{bmatrix}^{1}$$

1. The column matrix X can also be written $[x\ y]^T$ which is sometimes a more desirable form.

3. MOMENTS OF A DISTRIBUTION

The concept of moments is utilized for a geometrical interpretation of the parameters of an error distribution. The first moment of mass is the product of the mass and the moment arm. The first moment, when divided by the mass, is defined as the center of mass. The second moment, moment of inertia, is defined as the sum of the products obtained by multiplying each elementary mass by the square of its distance from the line. The second moment, when divided by the mass, is defined as the square of the radius of gyration. Considering a distribution representing a hypothetical unit mass, the moments of the hypothetical mass correspond to the statistical parameters of the distribution.

Consider the distribution of missile impacts around an aiming point, the performance minus computed values (impact minus aiming point) are expressed as downrange and crossrange errors, fig. 1.



The relationships between statistical parameters and moments are used to derive the distribution function [f(x,y)] which describes the frequency distribution of the errors. The errors are assumed to follow a normal distribution.

The analysis of a distribution by its moments, in the pure mathematical sense, is complicated. For convenience it is assumed that the moments, as found by succeeding sections, describe a unique distribution.[1]

3.1. First Moment. The first moment corresponds to the mean value of a distribution. The mean values $(\bar{x},\bar{y})$ of the crossrange and downrange errors are computed by:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \; ; \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

where:

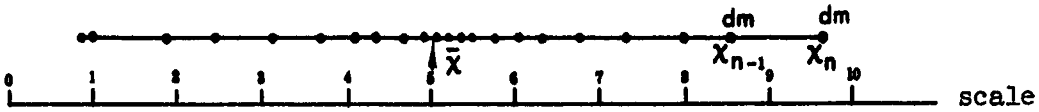
x = crossrange error

y = downrange error

The correspondence between the mean value and first moment is apparent from the definition of the first moment and can be visualized by considering small masses ($dm = 1$) with coordinates x_i along an assumed straight bar (whose weight is neglected). The first moment or

1. A thorough discussion concerning moments and uniquiness of distributions is contained in references Nr. 4 and 12.

center of mass of the system of particles (x_i) is:



$$\bar{x}_{dm} = \frac{\sum_{i=1}^{n} x_i}{n}$$

similarly

$$\bar{y}_{dm} = \frac{\sum_{i=1}^{n} y_i}{n}$$

All masses dm are considered equal to 1.

The values $\bar{x}$ and $\bar{y}$ are the centers of the two distributions f(x) and f(y), the point $(\bar{x},\bar{y})$ is then the center of the combined distribution f(x,y). For convenience of notation and computation the coordinate system is considered as having its origin at $(\bar{x},\bar{y})$.

In the theoretical approach, the mean value is the expected value. Since all possible values must be considered, the theoretical first moments are:

$$E\{x\} = \int_{-\infty}^{+\infty} xf(x)dx \equiv \mu_x$$

$$E\{y\} = \int_{-\infty}^{+\infty} yf(y)dy \equiv \mu_y$$

and:

$$\bar{x} \to \mu_x \text{ as } n \to \infty$$

$$\bar{y} \to \mu_y \text{ as } n \to \infty$$

3.2. Second Moment. The second moment, the moment of inertia, corresponds to the variance of a distribution. The second moment of an object describes a physical characteristic of the mass, the manner of which the mass is distributed within the body; the variance denotes the similar characteristic within a distribution. The second moment, about the coordinate axis, of the system of particles x_i, introduced in section 3.1. is:

$$\sum_{i=1}^{n} x_i^2 dm$$

thus the theoretical second moment is:

$$\int x^2 \partial m$$

The variance by definition is the mean of the squares of all the errors and computed by:

$$\sigma_x^2 = \frac{\sum_{i=1}^{n} x_i^2}{n}$$

where x is the error.

Written in theoretical sense:

$$\sigma_x^2 = \frac{\int_{-\infty}^{+\infty} x^2 f(x)dx}{\int_{-\infty}^{+\infty} f(x)dx}$$

since:

$$\int_{-\infty}^{+\infty} f(x)dx \equiv 1$$

$$\sigma_x^2 = \int_{-\infty}^{+\infty} x^2 f(x)dx = \mu_{xx} = E\left\{x^2\right\}$$

similarly:

$$\sigma_y^2 = \int_{-\infty}^{+\infty} y^2 f(y)dy = \mu_{yy} = E\left\{y^2\right\}$$

A standard error ellipse can now be drawn with its center at the origin and σ_x and σ_y as its semimajor and semiminor axes. The ellipse displays the concentration of the distribution of downrange and crossrange errors in the x and y directions.

3.3. Mixed Product Moment. The covariance is the mixed product moment. Although first and second moments of x and y were described earlier, when the moments of two or more variables are considered, a third type, the mixed product moment (covariance), exists. By definition the covariance is:

$$\mu_{xy} \equiv \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy \equiv E\{x,y\}$$

where: integration takes place over the entire area of the distribution.

3.4. Moment Generating Function. The moment generating function is introduced to avoid difficult integration which is encountered in moment (variance-covariance) functions. The moment generating function for the univariate case (single variable) is defined as:

$$E\left\{e^{\theta x}\right\} = \int_{-\infty}^{+\infty} e^{\theta x} f(x)dx$$

where: θ is a mathematical device introduced for the purpose of defining the moments.

Developing $e^{\theta x}$ in power series:

$$e^{\theta x} = 1 + \theta x + \frac{\theta^2 x^2}{2!} + \frac{\theta^3 x^3}{3!} + \ldots$$

Integrating the power series:

$$E\left\{e^{\theta x}\right\} = \int_{-\infty}^{+\infty} f(x)dx + \int_{-\infty}^{+\infty} \theta x f(x)dx + \int_{-\infty}^{+\infty} \frac{\theta^2 x^2}{2!} f(x)dx + \int_{-\infty}^{+\infty} \frac{\theta^3 x^3}{3!} f(x)dx + \ldots$$

$$= \int_{-\infty}^{+\infty} f(x)dx + \theta \int_{-\infty}^{+\infty} x f(x)dx + \frac{\theta^2}{2!} \int_{-\infty}^{+\infty} x^2 f(x)dx + \frac{\theta^3}{3!} \int_{-\infty}^{+\infty} x^3 f(x)dx + \ldots$$

$$\ldots + \frac{\theta^k}{k!} \int_{-\infty}^{+\infty} x^k f(x)dx + \ldots$$

From previous definitions:

$$\int_{-\infty}^{+\infty} f(x)dx = 1; \quad \int_{-\infty}^{+\infty} x f(x)dx = \mu_x; \quad \int_{-\infty}^{+\infty} x^2 f(x)dx = \mu_{xx}$$

and:

$$\int_{-\infty}^{+\infty} x^k f(x)dx \equiv k^{th} \text{ moment of } x$$

Therefore:

$$E\left\{e^{\theta x}\right\} = 1 + \theta\mu_x + \frac{\theta^2}{2!}\mu_{xx} + \cdots$$

The successive derivatives of $E\left\{e^{\theta x}\right\}$ with respect to θ, evaluated at $\theta = 0$ yields the corresponding moments.

Since:

$$\frac{d^k}{d\theta^k} e^{\theta x} = x^k e^{\theta x}$$

at $\theta = 0$

$$\int_{-\infty}^{+\infty} \frac{d^k}{d\theta^k} e^{\theta x} f(x)dx = \int_{-\infty}^{+\infty} x^k f(x)dx \equiv E\left\{x^k\right\}$$

Therefore the moment k of the distribution function may be obtained by the k^{th} derivative of the function. This property will be applied later.

3.5. Moment Matrix. The moment matrix (M) contains the parameters of the distribution and is used as a convenient notation describing the distribution function.

$$M = \begin{bmatrix} \mu_{xx} & \mu_{xy} \\ \mu_{yx} & \mu_{yy} \end{bmatrix}$$

where:

μ_{xx} = variance in $x(\sigma_x^2)$

μ_{yy} = variance in $y(\sigma_y^2)$

μ_{xy} = μ_{yx} = covariance of x and y

$$\rho = \frac{\mu_{xy}}{\sigma_x \sigma_y} \text{ (assumed value)}$$ [1]

ρ = correlation coefficient

let $\sigma_{xy} = \rho\sigma_x\sigma_y = \mu_{xy}$

and:

$$-1 \leqq \rho \leqq +1$$

M will always be a symmetric matrix.

The rank (r) of M displays the properties of a distribution.

When:

$r = 0$; $\mu_{xx} = \mu_{yy} = 0$ the mass is concentrated at a point, the center of mass.

$r = 1$; $\mu_{xx} \cdot \mu_{yy} = 0$ the mass is concentrated on a straight line passing through the center of mass.

$r > 1$; there is no straight line which contains the total mass of the distribution.

4. APPLICATION OF MOMENTS TO THE NORMAL ERROR DISTRIBUTION

The normal error distribution is written in the form:

$$f = K_2 \exp\left\{-\frac{1}{2} Q_2\right\}$$ [2]

where:

Q_2 is the quadratic expression of the error distribution.

K_2 is the constant corresponding to a particular Q_2.

---

1. To be proven later.
2. For notation purposes the exponential values (e^x) will be written $\exp\{x\}$.

The general normal bivariate error distribution form, expressed by its statistical parameters, is obtained by use of the moment generating function and moment matrix.

4.1. Determination of K_2 Constant for the Normal Bivariate Error Distribution. The quadratic expression Q_2 for the bivariate error distribution is written:

$$Q_2 = a_{11}X^2 + 2a_{12}XY + a_{22}Y^2$$

Then, from section 2.4.:

$$Q_2 = [X,Y] \; A \begin{bmatrix} X \\ Y \end{bmatrix} = [X,Y] \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

where:

$X = (x - \bar{x})$

$Y = (y - \bar{y})$

For convenience let $\bar{x} = \bar{y} = 0$

For a definite positive quadratic form it is necessary that:

$$a_{11} > 0$$
$$a_{12} > 0$$
$$a_{11}a_{22} > a_{12}^2$$

The fact that, over the entire x,y range, the integral of the f(x,y) function is unity (1) is used to determine K_2. That is:

$$1 = K_2 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\left(a_{11}x^2 + 2a_{12}xy + a_{22}y^2\right)\right\} dxdy$$

completing the square of Q_2:

$$a_{11}\left(x + \frac{a_{12}}{a_{11}}\, y\right)^2 + a_{22}y^2 - \frac{a_{12}^2}{a_{11}}\, y^2$$

$$a_{11}\left(x^2 + 2\,\frac{a_{12}}{a_{11}}\, xy + \frac{a_{12}^2}{a_{11}^2}\, y^2\right) + a_{22}y^2 - \frac{a_{12}^2}{a_{11}}\, y^2$$

Thus:

$$1 = K_2 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\, a_{11}\left(x^2 + 2\,\frac{a_{12}}{a_{11}}\, xy + \frac{a_{12}^2}{a_{11}^2}\, y^2\right) - \frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right) y^2\right\} dxdy$$

Letting:

$$Z_1 = \left(x + \frac{a_{12}}{a_{11}}\, y\right)$$

and using:

$$\int_{-\infty}^{+\infty} \exp\left\{-\, cx^2\right\} dx = \sqrt{\frac{\pi}{c}}$$

Then:

$$1 = K_2 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\, a_{11}\left(Z_1\right)^2 - \frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right) y^2\right\} dZ_1 dy$$

Letting:

$$c_1 = \frac{1}{2}\, a_{11} \qquad \text{and} \qquad c_2 = \frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)$$

$$1 = K_2 \int_{-\infty}^{+\infty} \exp\left\{-\, c_1 Z_1^2\right\} dZ \cdot \int_{-\infty}^{+\infty} \exp\left\{-\, c_2 y^2\right\} dy$$

$$1 = K_2 \sqrt{\frac{\pi}{\frac{1}{2} a_{11}}} \cdot \sqrt{\frac{\pi}{\frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)}}$$

$$= K_2 \frac{2\pi}{\sqrt{a_{11}a_{22} - a_{12}^2}}$$

$$K_2 = \frac{\sqrt{a_{11}a_{22} - a_{12}^2}}{2\pi}$$

The factor $a_{11}a_{22} - a_{12}^2$ is the determinant of matrix A. Therefore:

$$K_2 = \frac{\sqrt{|A|}}{2\pi}$$

and the general bivariate normal error distribution is written:

$$f(x,y) = \frac{\sqrt{|A|}}{2\pi} \exp\left\{- \frac{1}{2}\left(a_{11}x^2 + 2a_{12}xy + a_{22}y^2\right)\right\}$$

$$= \frac{\sqrt{|A|}}{2\pi} \exp \; - \frac{1}{2}\left\{Q_2\right\}$$

4.2. <u>The Bivariate Error Distribution in Terms of Statistical Parameters</u>. To determine the significance of the elements a_{11}, a_{22} and a_{12} of $f(x,y)$, assume a normal bivariate distribution and determine its variance and covariance by use of the moment generating function. From section 3.4. the moment generating function for the univariate case is:

$$E\left\{e^{\theta x}\right\} = \int_{-\infty}^{+\infty} e^{\theta x} f(x)dx$$

For the bivariate case:

$$E\left\{\exp\,[\theta x + \theta y]\right\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{\theta x + \theta y\right\} f(x,y)\,dxdy$$

where:

$$f(x,y) = \frac{\sqrt{|A|}}{2\pi} \exp\left\{-\frac{1}{2}\,Q_2\right\}$$

Therefore:

$$E\left\{\exp\,[\theta x + \theta y]\right\} = \frac{\sqrt{|A|}}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\,Q_2 + \theta x + \theta y\right\}dxdy$$

or:

$$E\left\{\exp\,[\theta x + \theta y]\right\} = \frac{\sqrt{|A|}}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\left(a_{11}x^2 + 2a_{12}xy + a_{22}y^2\right) + \theta_x x + \theta_y y\right\}dxdy$$

Expanding the right side of this equation:[1]

$$E\left\{\exp\,[\theta x + \theta y]\right\} = \frac{\sqrt{|A|}}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\,a_{11}\left(x + \frac{a_{12}}{a_{11}}\,y - \frac{\theta_x}{a_{11}}\right)^2 - \frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)\left(y + \frac{a_{12}\theta_x - a_{11}\theta_y}{|A|}\right)^2 + \frac{1}{2}\,R\right\}dxdy$$

where:

$$R = \frac{a_{22}\theta_x^2 + a_{11}\theta_y^2 - 2a_{12}\theta_x\theta_y}{|A|} = a^{11}\theta_x^2 + a^{22}\theta_y^2 + 2a^{12}\theta_x\theta_y$$

---

1. The details of this expansion are shown in Appendix A

Letting:

$$Z_1 = \left(x + \frac{a_{12}}{a_{11}} y - \frac{\theta_x}{a_{11}}\right); \qquad Z_2 = \left(y + \frac{a_{12}\theta_x - a_{11}\theta_y}{|A|}\right)$$

Then:

$$E\left\{\exp\left[\theta x + \theta y\right]\right\} = \frac{\sqrt{|A|}}{2\pi} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}a_{11}Z_1^2\right\} dZ_1 \cdot \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right) Z_2^2\right\} dZ_2$$

$$\cdot \exp\left\{+\frac{1}{2}\left(a^{11}\theta_x^2 + a^{22}\theta_y^2 + 2a^{12}\theta_x\theta_y\right)\right\}$$

Letting:

$$C_1 = \frac{1}{2}a_{11}, \qquad C_2 = \frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)$$

$$E\left\{\exp\left[\theta x + \theta y\right]\right\} = \frac{\sqrt{|A|}}{2\pi} \int_{-\infty}^{+\infty} \exp\left\{-C_1Z_1^2\right\} dZ_1 \cdot \int_{-\infty}^{+\infty} \exp\left\{-C_2Z_2^2\right\} dZ_2$$

$$\cdot \exp\left\{+\frac{1}{2}\left(a^{11}\theta_x^2 + a^{22}\theta_y^2 + 2a^{12}\theta_x\theta_y\right)\right\}$$

$$= \frac{\sqrt{|A|}}{2\pi} \cdot \frac{\sqrt{\pi}}{\sqrt{\frac{1}{2}a_{11}}} \cdot \frac{\sqrt{\pi}}{\sqrt{\frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)}}$$

$$\cdot \exp\left\{+\frac{1}{2}\left(a^{11}\theta_x^2 + a^{22}\theta_y^2 + 2a^{12}\theta_x\theta_y\right)\right\}$$

For convenience, a^{11}, a^{22}, and a^{12} no longer denote cofactors (p 10) but denote the elements $\frac{a_{22}}{|A|}$, $\frac{a_{11}}{|A|}$, and $\frac{a_{12}}{|A|}$ respectively.

$$= \frac{\sqrt{|A|}}{2} \cdot \frac{2}{\sqrt{a_{11}a_{22} - a_{12}^2}}$$

$$\cdot \exp\left\{+\frac{1}{2}\left(a^{11}\theta_x^2 + a^{22}\theta_y^2 + 2a^{12}\theta_x\theta_y\right)\right\}$$

Therefore:

$$E\left\{\exp\left[\theta x + \theta y\right]\right\} = \exp\left\{+\frac{1}{2}\left(a^{11}\theta_x^2 + a^{22}\theta_y^2 + 2a^{12}\theta_x\theta_y\right)\right\}$$

The variances σ_x^2 and σ_y^2 are obtained by evaluating the second partial derivatives of $E\left\{\exp\left[\theta x + \theta y\right]\right\}$ with respect to θ_x and θ_y respectively at $\theta_x = \theta_y = 0$. For convenience let E denote $E\left\{\exp\left[\theta x + \theta y\right]\right\}$ in the following presentation:

$$\sigma_x^2 = \frac{\partial^2 E}{\partial \theta_x^2}$$

$$\frac{\partial E}{\partial \theta_x} = E\left(a^{11}\theta_x + a^{12}\theta_y\right)$$

$$\frac{\partial^2 E}{\partial \theta_x^2} = E \cdot a^{11} + \left(a^{11}\theta_x + a^{12}\theta_y\right)\left\{E\left(a^{11}\theta_x + a^{12}\theta_y\right)\right\}$$

At $\theta_x = \theta_y = 0$

$$\sigma_x^2 = \frac{\partial^2 E}{\partial \theta_x^2} = a^{11}$$

Similarly:

$$\sigma_y^2 = \frac{\partial^2 E}{\partial \theta_y^2}$$

$$\frac{\partial E}{\partial \theta_y} = E\left(a^{22}\theta_y + a^{12}\theta_x\right)$$

$$\frac{\partial^2 E}{\partial \theta_y^2} = E \cdot a^{22} + \left(a^{22}\theta_y + a^{12}\theta_x\right)\left\{E\left(a^{22}\theta_y + a^{12}\theta_x\right)\right\}$$

At $\theta_x = \theta_y = 0$

$$\sigma_y^2 = \frac{\partial^2 E}{\partial \theta_y^2} = a^{22}$$

The covariance

$$\sigma_{xy} = E\{(x - \bar{x})(y - \bar{y})\} = \frac{\partial^2 E}{\partial \theta_x \partial \theta_y}$$

$$\frac{\partial E}{\partial \theta_y} = E\left(a^{22}\theta_y + a^{12}\theta_x\right)$$

$$\frac{\partial^2 E}{\partial \theta_x \partial \theta_y} = E \cdot a^{12} + \left(a^{22}\theta_y + a^{12}\theta_x\right)\left\{E\left(a^{11}\theta_x + a^{12}\theta_y\right)\right\}$$

$$= E\left\{a^{12} + \left(a^{22}\theta_y + a^{12}\theta_x\right)\left(a^{11}\theta_x + a^{12}\theta_y\right)\right\}$$

At $\theta_x = \theta_y = 0$

$$\sigma_{xy} = \frac{\partial^2 E}{\partial \theta_x \partial \theta_y} = a^{12}$$

From the previous discussion:

σ_x^2 = second moment of x

σ_y^2 = second moment of y

$\sigma_{xy} = \rho\sigma_x\sigma_y$ = mixed product moment of x and y but:

$$a^{11} = \sigma_x^2 = \frac{a_{22}}{|A|}$$

$$a^{22} = \sigma_y^2 = \frac{a_{11}}{|A|}$$

$$a^{12} = \rho\sigma_x\sigma_y = \frac{a_{12}}{|A|}$$

Therefore the moment matrix of the bivariate normal distribution is:

$$M = \begin{bmatrix} a^{11} & a^{12} \\ a^{12} & a^{22} \end{bmatrix} = A^{-1} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$$

Equivalently $A = M^{-1}$

Following the procedure outlined in paragraph 2.2. for inversion:

$$M^{-1} = \begin{bmatrix} \dfrac{\sigma_y^2}{\sigma_x^2\sigma_y^2 - \rho^2\sigma_x^2\sigma_y^2} & \dfrac{-\rho\sigma_x\sigma_y}{\sigma_x^2\sigma_y^2 - \rho^2\sigma_x^2\sigma_y^2} \\[2ex] \dfrac{-\rho}{\sigma_x^2\sigma_y^2 - \rho^2\sigma_x^2\sigma_y^2} & \dfrac{\sigma_x^2}{\sigma_x^2\sigma_y^2 - \rho^2\sigma_x^2\sigma_y^2} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{1}{\sigma_x^2(1-\rho^2)} & \dfrac{-\rho}{\sigma_x\sigma_y(1-\rho^2)} \\[2ex] \dfrac{-\rho}{\sigma_x\sigma_y(1-\rho^2)} & \dfrac{1}{\sigma_y^2(1-\rho^2)} \end{bmatrix}$$

Thus:

$$K_2 = \frac{\sqrt{|A|}}{2\pi} = \frac{\sqrt{|M^{-1}|}}{2\pi}$$

$$K_2 = \frac{1}{2\pi}\sqrt{\frac{1}{\sigma_x^2(1-\rho^2)}\cdot\frac{1}{\sigma_y^2(1-\rho^2)} - \frac{\rho^2}{\sigma_x^2\sigma_y^2(1-\rho^2)^2}}$$

$$= \frac{1}{2\pi}\sqrt{\frac{1}{\sigma_x^2\sigma_y^2(1-\rho^2)^2} - \frac{\rho^2}{\sigma_x^2\sigma_y^2(1-\rho^2)^2}}$$

$$= \frac{1}{2\pi}\sqrt{\frac{(1-\rho^2)}{\sigma_x^2\sigma_y^2(1-\rho^2)^2}}$$

$$K_2 = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$$

From 4.1.

$$Q_2 = [X\ Y]\ A\ [X\ Y]^T = [X\ Y]\ M^{-1}\ [X\ Y]^T$$

$$= [X\ Y] \begin{bmatrix} \frac{1}{\sigma_x^2(1-\rho^2)} & \frac{-\rho}{\sigma_x\sigma_y(1-\rho^2)} \\ \frac{-\rho}{\sigma_x\sigma_y(1-\rho^2)} & \frac{1}{\sigma_y^2(1-\rho^2)} \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

$$= \frac{X^2}{\sigma_x^2(1-\rho^2)} + \frac{Y^2}{\sigma_y^2(1-\rho^2)} - \frac{2\rho XY}{\sigma_x\sigma_y(1-\rho^2)}$$

Thus $f(X,Y) = K_2 \exp\left\{-\frac{1}{2}Q_2\right\}$

$$f(X,Y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{X^2}{\sigma_x^2} + \frac{Y^2}{\sigma_y^2} - \frac{2\rho XY}{\sigma_x\sigma_y}\right)\right\}$$

Since:

$$X = (x - \bar{x}) \quad \text{and} \quad Y = (y - \bar{y})$$

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2} - \frac{2\rho\,(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y}\right)\right\}^{1}$$

If $\rho = 0$ (case of independent errors)

$$M = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}$$

and:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left(\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2}\right)\right\}$$

Showing that independent errors are merely a special case of dependent errors.

5. ANALYSIS OF THE NORMAL BIVARIATE ERROR DISTRIBUTION

In section 4, the equations of the normal bivariate error distribution were determined, and the elements of M were shown as variances and

---

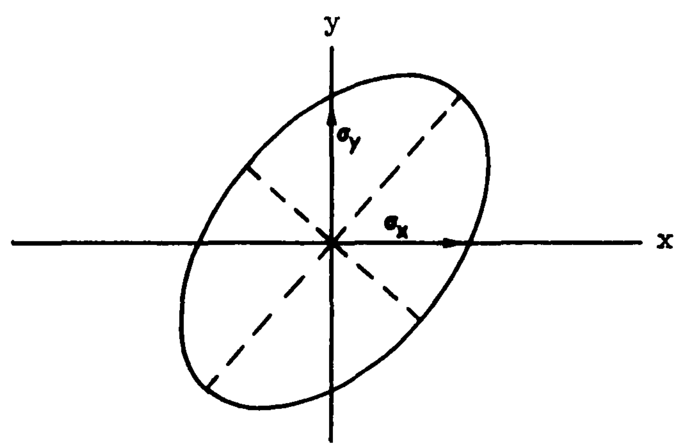
1. This equation results from the assumption $\rho = \frac{\mu_{xy}}{\sigma_x\sigma_y}$ introduced in section 3.5. The assumption may now be proven correct; Appendix C.

covariances. The covariances for independent errors were zero. This section is concerned with reducing a normal error distribution with dependent errors to a form expressible as independent errors (covariances zero). After this is accomplished, circular errors are obtained by use of the methods presented in ACIC TR-96.

The quadratic expression for independent (Q_2^I) and dependent (Q_2^D) bivariate error distributions are:

$$Q_2^I = \left(\frac{X^2}{\sigma_x^2} + \frac{Y^2}{\sigma_y^2}\right); \qquad Q_2^D = \frac{1}{1-\rho^2}\left(\frac{X^2}{\sigma_x^2} + \frac{Y^2}{\sigma_y^2} - \frac{2\rho XY}{\sigma_x\sigma_y}\right)$$

When covariances are zero the major and minor axes of the ellipse coincide with the axes of the coordinate system. When the covariances are not zero, there are dependent errors and the axes of the error ellipse do not coincide with the coordinate system. Therefore, removal of the non-zero covariances from dependent errors is equivalent to a rotation of axes.



The process of diagonalization of the covariance matrix is used for the rotation.

Diagonalization is interpreted as follows: If a point p is allowed to move freely around the perimeter of an ellipse, there are four positions where the rate of change of the distance (D^2) between these positions and the center of the ellipse is zero. These four positions correspond to the points of intersection of the principal axes of the ellipse and its perimeter. The problem then is to find the maximum and minimum value of p as it traverses the elliptical path. This involves finding the extreme values of a function of two variables with one side condition. Introducing the Lagrangian multiplier (λ):

$$D^2 = F + \lambda g$$

where:

F is $(x^2 + y^2)$ the square of the length of the radius vector, p.

g is the side condition (the equation of the ellipse).

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 - C = 0$$

or:

$$C - (a_{11}x^2 + 2a_{12}xy + a_{22}y^2) = 0$$

Thus:

$$D^2 = x^2 + y^2 + \lambda (C - a_{11}x^2 - 2a_{12}xy - a_{22}y^2)$$

The extreme values of D^2 occur when $\frac{\partial D^2}{\partial x}$ and $\frac{\partial D^2}{\partial y}$ equal zero.

$$\frac{\partial D^2}{\partial x} = 2x - \lambda\,(2a_{11}x + 2a_{12}y) = 0$$

$$= x - \lambda\,(a_{11}x + a_{12}y) = 0$$

similarly:

$$\frac{\partial D^2}{\partial y} = y - \lambda\,(a_{12}x + a_{22}y) = 0$$

Note that:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} a^{11} & a^{12} \\ a^{12} & a^{22} \end{bmatrix} = \begin{bmatrix} (a_{11}a^{11} + a_{12}a^{12} = 1), & (a_{11}a^{12} + a_{12}a^{22} = 0) \\ (a_{12}a^{11} + a_{22}a^{12} = 0), & (a_{12}a^{12} + a_{22}a^{22} = 1) \end{bmatrix}$$

Since $A \cdot A^{-1} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Multiplying the $\frac{\partial D^2}{\partial x}$ by a^{11} and $\frac{\partial D^2}{\partial y}$ by a^{12} yields:

$$a^{11}x - \lambda\,(a_{11}a^{11}x + a_{12}a^{11}y) = 0$$
$$a^{12}y - \lambda\,(a_{12}a^{12}x + a_{22}a^{12}y) = 0$$

Adding:

$$a^{11}x + a^{12}y - \lambda \left\{ x\,(a_{11}a^{11} + a_{12}a^{12}) + y\,(a_{12}a^{11} + a_{22}a^{12}) \right\} = 0$$

or:

$$a^{11}x + a^{12}y - \lambda x = 0$$

Similarly, multiplying $\frac{\partial D^2}{\partial x}$ by a^{12} and $\frac{\partial D^2}{\partial y}$ by a^{22} yields:

$$a^{12}x - \lambda\,(a_{11}a^{12}x + a_{12}a^{12}y) = 0$$
$$a^{22}y - \lambda\,(a_{12}a^{22}x + a_{22}a^{22}y) = 0$$

Adding:

$$a^{12}x + a^{22}y - \lambda\left\{x\,(a_{11}a^{12} + a_{12}a^{22}) + y\,(a_{12}a^{12} + a_{22}a^{22})\right\} = 0$$

or:

$$a^{12}x + a^{22}y - \lambda y = 0$$

Therefore:

$$a^{11}x + a^{12}y - \lambda x = 0$$
$$a^{12}x + a^{22}y - \lambda y = 0$$

Collecting terms:

$$x\,(a^{11} - \lambda) + a^{12}y = 0$$
$$xa^{12} + (a^{22} - \lambda)\,y = 0$$

To have non-zero values for x and y the determinant of the coefficients of the two equations must equal zero.

$$\begin{vmatrix} a^{11} - \lambda & a^{12} \\ a^{12} & a^{22} - \lambda \end{vmatrix} = 0$$

The expansion of the determinant is considered the characteristic equation which, in the bivariate case, is:

$$\lambda^2 - (a^{11} + a^{22})\lambda + a^{11}a^{22} - (a^{12})^2 = 0$$

Solution by the quadratic formula yields:

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where:

$$a = 1$$

$$b = -(a^{11} + a^{22}) = -(\sigma_x^2 + \sigma_y^2)$$

$$c = a^{11}a^{22} - (a_{12})^2 = \sigma_x^2\sigma_y^2(1 - \rho^2)$$

Therefore:

$$\lambda = \frac{+(\sigma_x^2 + \sigma_y^2) \pm \sqrt{(\sigma_x^2 + \sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2(1 - \rho^2)}}{2}$$

$$\lambda = \frac{(\sigma_x^2 + \sigma_y^2)}{2} \pm \frac{\sqrt{(\sigma_x^4 - 2\sigma_x^2\sigma_y^2 + \sigma_y^4) + 4\sigma_x^2\sigma_y^2\rho^2}}{2}$$

$$\lambda = \frac{(\sigma_x^2 + \sigma_y^2)}{2} \pm \frac{\sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\sigma_x^2\sigma_y^2\rho^2}}{2}$$

But for independent errors $\rho = 0$.

Consequently:

$$\lambda = \frac{(\sigma_x^2 + \sigma_y^2)}{2} \pm \frac{\sqrt{(\sigma_x^2 - \sigma_y^2)^2}}{2}$$

or:

$$\lambda = \frac{(\sigma_x^2 + \sigma_y^2)}{2} \pm \frac{(\sigma_x^2 - \sigma_y^2)}{2}$$

$$\lambda_1 = \frac{(\sigma_x^2 + \sigma_y^2)}{2} + \frac{(\sigma_x^2 - \sigma_y^2)}{2} = \sigma_x^2$$

$$\lambda_2 = \frac{(\sigma_x^2 + \sigma_y^2)}{2} - \frac{(\sigma_x^2 - \sigma_y^2)}{2} = \sigma_y^2$$

That is, the square roots of λ_1, and λ_2 correspond to the maximum and minimum values of D^2. Since λ_1, and λ_2 also correspond to the squares of the lengths of the principal axes of the standard error ellipse, it can be concluded that the magnitudes of the principal axes are $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$. This is true for dependent errors as well as for independent errors. The process of diagonalization reduces to the formation and solution of the characteristic equation of the determinant of the inverse of the moment matrix.

6. ANALYSIS OF THE NORMAL TRIVARIATE ERROR DISTRIBUTION

The theory of the normal bivariate error distribution can be generalized to cover the normal error distribution for 3 or more variables. In this section the trivariate case is considered, but a complete derivation is not presented because the end value is visualized as an extension of the bivariate case.

6.1. General Form. The general normal trivariate error distribution is written:

$$f(x,y,z) = K_3 \exp\left\{-\frac{1}{2} Q_3\right\}$$

where:

$$K_3 = \frac{\sqrt{A}}{\sqrt{(2\pi)^3}} \quad [1]$$

$$Q_3 = a_{11}X^2 + 2a_{12}XY + 2a_{13}XZ + a_{22}Y^2 + 2a_{23}YZ + a_{33}Z^2$$

$$X = (x - \bar{x}) \qquad Y = (y - \bar{y}) \qquad Z = (z - \bar{z})$$

For convenience let $\bar{x} = \bar{y} = \bar{z} = 0$

The moments of the distribution are shown as:

$$\begin{array}{lrcccc}
\text{variance of} & x & = & a^{11} & = & \sigma_x^2 \\
 & y & = & a^{22} & = & \sigma_y^2 \\
 & z & = & a^{33} & = & \sigma_z^2 \\
\text{covariance of} & x \text{ and } y & = & a^{12} & = & \sigma_{xy} \\
 & x \text{ and } z & = & a^{13} & = & \sigma_{xz} \\
 & y \text{ and } z & = & a^{23} & = & \sigma_{yz}
\end{array}$$

6.2. Axes of the Error Ellipsoid. The characteristic equation for the trivariate case is obtained by expanding the determinant

$$\begin{vmatrix} a^{11} - \lambda & a^{12} & a^{13} \\ a^{12} & a^{22} - \lambda & a^{23} \\ a^{13} & a^{23} & a^{33} - \lambda \end{vmatrix} = 0$$

---

1. Derived in Appendix B

It will be of the form:

$$\lambda^3 - r\lambda^2 + s\lambda - t = 0$$

where:

$$r = a^{11} + a^{22} + a^{33}$$
$$s = a^{11}a^{22} + a^{11}a^{33} + a^{22}a^{33} - (a^{12})^2 - (a^{13})^2 - (a^{23})^2$$
$$t = a^{11}a^{22}a^{33} + 2a^{12}a^{13}a^{23} - a^{11}(a^{23})^2 - a^{22}(a^{13})^2 - a^{33}(a^{12})^2$$

The first root (λ_a) is found by applying Newton's method to the polynomial in question ($\lambda^3 - r\lambda^2 + s\lambda - t = 0$) and utilizing the first derivative and an approximate value for λ_a. By successive iteration the root can be solved for as many correct digits as required. After the value of one root is obtained, the cubic can be reduced to a second degree polynomial and solved by the quadratic formula for the remaining two roots. The procedure is as follows:

$$f(\lambda) = \lambda^3 - r\lambda^2 + s\lambda - t = 0$$

$$f'(\lambda) = 3\lambda^2 - 2r\lambda + t$$

Solution of first root:

Let λ_1, be first approximation[1]

$$\lambda_2 = \lambda_1 - \frac{f(\lambda_1)}{f'(\lambda_1)}$$

1. A first approximation is the larger of the diagonal elements, a^{11}, a^{22} or a^{33}.

successive iteration:

$$\lambda_3 = \lambda_2 - \frac{f(\lambda_2)}{f'(\lambda_2)}$$

$$\cdots\cdots\cdots\cdots$$

$$\lambda_n = \lambda_{n-1} - \frac{f(\lambda_{n-1})}{f'(\lambda_{n-1})}$$

Let:

$$\lambda_a = \lambda_n$$

Applying synthetic division for the remaining roots:

$$a\lambda^2 + (b + a\lambda_a)\,\lambda + \left\{ c + (b + a\lambda_a)\,\lambda_a \right\} = 0$$

Let:

$$A = a$$

$$B = b + a\lambda_a$$

$$C = c + (b + a\lambda_a)\,\lambda_a$$

Then:

$$\lambda_b = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

$$\lambda_c = \frac{-B - \sqrt{B^2 - 4AC}}{2A}$$

Therefore, the variances are known and the axes of standard error ellipsoid are:

$$a = \sqrt{\lambda_a}$$

$$b = \sqrt{\lambda_b}$$

$$c = \sqrt{\lambda_c}$$

7. ORIENTATION OF THE ERROR ELLIPSE AND ELLIPSOID

The orientation of the error ellipse or ellipsoid must be determined to complete the geometrical description of an error distribution. It has been concluded in previous sections that the covariance matrix contains the constant coefficients of the quadratic expression of the error ellipse and that the principal axes of the error ellipse are obtained by diagonalization of the covariance matrix. The operation provides the maximum and minimum errors of prescribed probabilities and is comparable to expressing the error ellipse in its standard form. It is possible to construct equi-probability curves around the mean value of the error distribution by multiplying these independent error values by a particular conversion factor (Theory of Errors, ACIC TR-96). To construct the equi-probability curves, the orientation with respect to the original axes must be determined by solving for the direction cosines between the original coordinate axes and the principal axes of the error ellipse. The following general theory and solution holds for both the bivariate and trivariate cases.

7.1. General Theory. Matrix theory coincides with and makes explicit use of vectors and vector properties. The rows of a matrix are considered row vectors and the columns, column vectors. The elements of the matrix are components of the vectors. The solution for the direction cosines involves determination of the cosine of the angle between the original coordinate axes, (x,y), and the positive direction of the principal axes, (x',y') of the error ellipse. From vector analysis:

$$\vec{a} \cdot \vec{b} = |\vec{a}|\ |\vec{b}| \cos \alpha \qquad (1)$$

where:

$\vec{a} \cdot \vec{b}$ = dot product of the two vectors a and b

$|\vec{a}|\ |\vec{b}|$ = product of the magnitude of the vectors a and b

$\cos \alpha$ = cosine of the angle between the vectors a and b

Diagonalization of the covariance matrix yielded values which were the latent roots (λ_1, λ_2, and λ_3) and corresponded to the squares of the lengths of the principal axes of the standard error ellipse. In converting the quadratic form of the dependent errors (covariance matrix with non-diagonal elements unequal to zero) to a quadratic form of independent errors (diagonal matrix), a coordinate transformation[1] has taken place. Therefore, there should be a transformation matrix which, when applied to the covariance matrix, will yield another matrix

---

1. In this instance, where the mean value of the distribution was chosen as the origin of the coordinate system, a rotation has taken place whereas in the general case it would be a transformation.

with the latent roots as its diagonal elements.

By definition,[1] the latent vector of a linear transformation, as described above, is:

$$AX = \lambda X \tag{2}$$

where:

A = covariance matrix

λ = latent root

X = latent vector

7.2. Direction Cosines for the Bivariate Case.

7.2.1. Derivation.

From eq. (2) $AX = \lambda X$

$$\begin{aligned} a_{11}X_1 + a_{12}Y_1 &= \lambda_1 X_1 \\ a_{12}X_1 + a_{22}Y_1 &= \lambda_1 Y_1 \end{aligned} \tag{3a}$$

or:

$$(a_{11} - \lambda_1)\, X_1 + a_{12}Y_1 = 0 \tag{3b1}$$

$$a_{12}X_1 + (a_{22} - \lambda_1)\, Y_1 = 0 \tag{3b2}$$

Since the cosines of the angles between the vectors are involved, it is possible to assign any value to one of the vector components and solve for the other with respect to the first.

Letting $X_1 = 1$ in equation (3b1)

---

1. Refn. Computational Methods of Linear Algebra, pp 37

Then:

$$Y_1 = \frac{-(a_{11} - \lambda_1)}{a_{12}}$$

or solving (3b2)

$$Y_1 = \frac{-a_{12}}{(a_{22} - \lambda)}$$

The values X_1 and Y_1 when placed in eq. 3b1 and 3b2 satisfy the equations and are, therefore, the latent vector components of the latent root (λ_1).

Similarly for the latent root (λ_2)

$$\begin{aligned} a_{11}X_2 + a_{12}Y_2 &= \lambda_2 X_2 \\ a_{12}X_2 + a_{22}Y_2 &= \lambda_2 Y_2 \end{aligned} \tag{4a}$$

or:

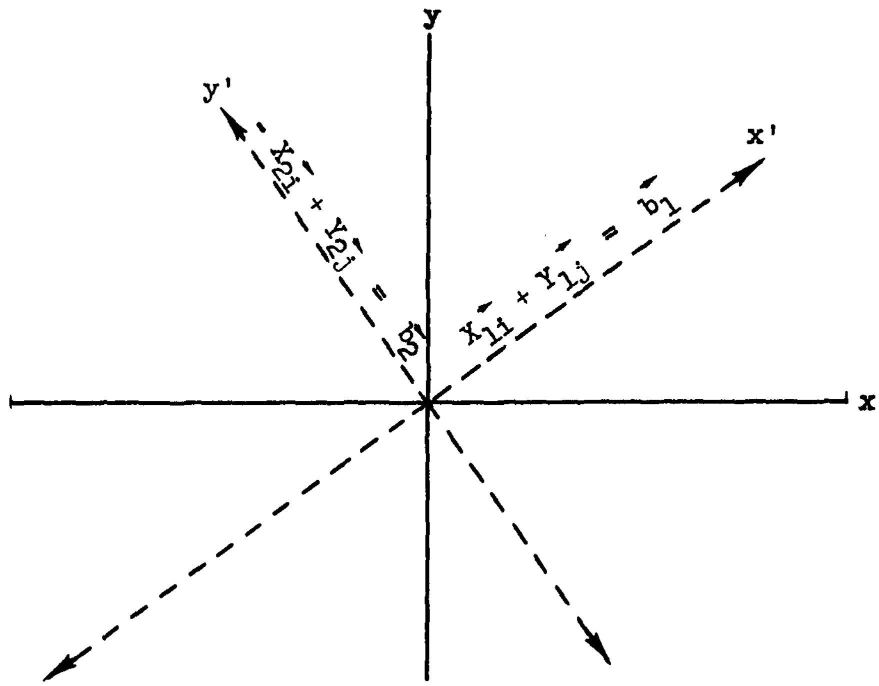
$$\begin{aligned} (a_{11} - \lambda_2)\, X_2 + a_{12}Y_2 &= 0 \\ a_{12}X_2 + (a_{22} - \lambda_2)\, Y_2 &= 0 \end{aligned}$$

Letting $X_2 = 1$

$$Y_2 = \frac{-a_{12}}{(a_{22} - \lambda_2)} = \frac{-(a_{11} - \lambda_2)}{a_{12}}$$

X_1, Y_1, X_2 and Y_2 are the vector components corresponding to the respective roots and are the components of the X matrix in equation (2).

$$\begin{bmatrix} X_1 & X_2 \\ Y_1 & Y_2 \end{bmatrix} \tag{5}$$



From eq. (1) ($\vec{a} \cdot \vec{b} = |\vec{a}|\ |\vec{b}| \cos \alpha$) it is necessary to compute the dot products of $\vec{a}$ and $\vec{b}$ and substitute them into eq. (1). The cosine of the angle between the +x and +x' axis is:

$$\vec{1}_i = \vec{a}_1 \text{ (unit vector along x axis)}$$

$$\vec{X}_{1i} + \vec{Y}_{1j} = \vec{b}_1 \text{ (vector along x' axis)}$$

$$\vec{a} \cdot \vec{b} = X_1$$

$$|\vec{a}|\ |\vec{b}| = \sqrt{1^2}\ \sqrt{X_1^2 + Y_1^2}$$

$$\therefore \cos \alpha_{11} = \frac{X_1}{\sqrt{X_1^2 + Y_1^2}}$$

Cosine of the angle between the +x and +y' axis:

$$\vec{1_i} = \vec{a_1}$$

$$-\vec{X_{2i}} + \vec{Y_{2j}} = \vec{b_2} \text{ (vector along +y' axis)}$$

$$\cos \alpha_{12} = \frac{-X_2}{\sqrt{X_2^2 + Y_2^2}}$$

Angle between the +y axis and +x' axis:

$$\vec{1_j} = \vec{a_2} \text{ (unit vector along +y axis)}$$

$$\vec{X_{1i}} + \vec{Y_{1j}} = \vec{b_1}$$

$$\cos \alpha_{21} = \frac{Y_1}{\sqrt{X_1^2 + Y_1^2}}$$

Angle between the +y axis and the +y' axis:

$$\vec{1_j} = \vec{a_2}$$

$$-\vec{X_{2i}} + \vec{Y_{2j}} = \vec{b_2}$$

$$\cos \alpha_{22} = \frac{Y_2}{\sqrt{X_2^2 + Y_2^2}}$$

7.2.2. Computation Method. The direction cosines are obtained by writing the latent vectors in form of eq. (5); summing the squares of all elements of each individual column; extracting the square roots; and dividing each element of the matrix (5) by the square root of its column.

$$
\begin{array}{c|ll}
 & x' & y' \\
\hline
x & \cos \alpha_{11} = \dfrac{X_1}{\sqrt{X_1^2 + Y_1^2}} , & \cos \alpha_{12} = \dfrac{- X_2}{\sqrt{X_2^2 + Y_2^2}} \\
y & \cos \alpha_{21} = \dfrac{Y_1}{\sqrt{X_1^2 + Y_1^2}} , & \cos a_{22} = \dfrac{Y_2}{\sqrt{X_2^2 + Y_2^2}}
\end{array}
\qquad (6)
$$

By definition the direction cosines measure the angle between the original positive axis and the positive principal axis of the error ellipse. The angles whose cosines are positive (+) are less than 90°; the angles whose cosines are negative ($- \alpha_{ij}$) are angles of 180° minus the angle of $+ \alpha_{ij}$.

For example consider the cosines as:

$$
\begin{array}{c|cc}
 & x' & y' \\
\hline
x & + & + \\
y & - & +
\end{array}
$$

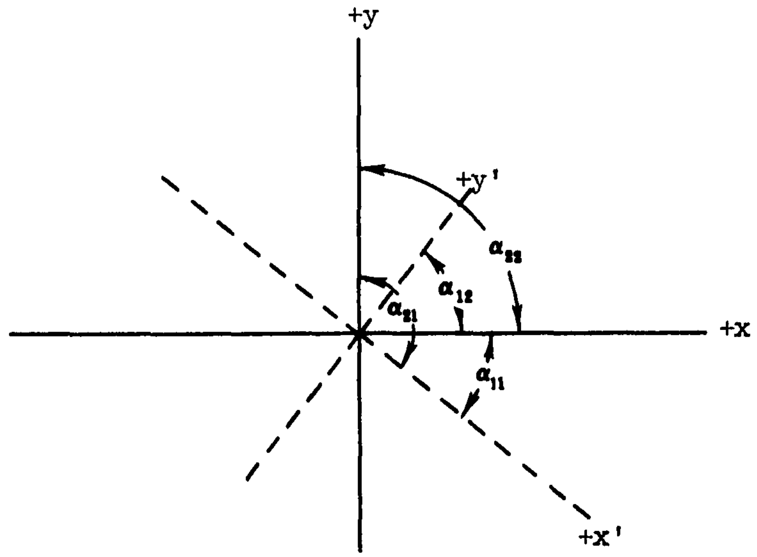
The angle between x and x' is 90° or less

y and x' is 90° or greater

Therefore, x' must be in the 4^{th} quadrant.

The angle between x and y' is 90° or less

y and y' is 90° or less



7.3. Direction Cosines for the Trivariate Case. The direction cosines for the trivariate case are found in the same manner as the bivariate case. However, the addition of a new coordinate axis (z) and another vector causes difficulty in the solution of the latent vectors. In the trivariate case there are 3 latent roots (λ) which must be solved to obtain 9 latent vector components.

From eq. (2): $AX = \lambda X$

For the latent root λ_n:

$$
\begin{aligned}
a_{11}X_n + a_{12}Y_n + a_{13}Z_n &= \lambda_n X_n \\
a_{12}X_n + a_{22}Y_n + a_{23}Z_n &= \lambda_n Y_n \\
a_{13}X_n + a_{23}Y_n + a_{33}Z_n &= \lambda_n Z_n
\end{aligned}
\tag{7a}
$$

Where n refers to any one of the 3 latent roots (λ_1, λ_2 or λ_3)

or:

$$
\begin{aligned}
(a_{11} - \lambda_n) X_n + a_{12}Y_n + a_{13}Z_n &= 0 \\
a_{12}X_n + (a_{22} - \lambda_n) Y_n + a_{23}Z_n &= 0 \qquad (7b) \\
a_{13}X_n + a_{23}Y_n + (a_{33} - \lambda_n) Z_n &= 0
\end{aligned}
$$

Since the determinant of this system equals zero[1] the solution presented in Coordinate Geometry, pp. 114-115 Theorem [22.1], is applicable.

Then:

$$X_n : Y_n : Z_n : = \left\{a_{12}a_{23} - (a_{22} - \lambda_n)\, a_{13}\right\} : - \left\{(a_{11} - \lambda_n)\, a_{23} - a_{12}a_{13}\right\} :$$

$$\left\{(a_{11} - \lambda_n)(a_{22} - \lambda_n) - a_{12}a_{12}\right\}$$

$$= \left\{(a_{22} - \lambda_n)(a_{33} - \lambda_n) - a_{23}a_{23}\right\} : - \left\{a_{12}\,(a_{33} - \lambda_n) - a_{13}a_{23}\right\} :$$

$$\left\{a_{12}a_{23} - a_{13}\,(a_{22} - \lambda_n)\right\} \qquad (8)$$

$$= \; - \left\{a_{12}\,(a_{33} - \lambda_n) - a_{23}a_{13}\right\} : \left\{(a_{11} - \lambda_n)(a_{33} - \lambda_n) - a_{13}a_{13}\right\} :$$

$$- \left\{(a_{11} - \lambda_n)\, a_{23} - a_{12}a_{13}\right\}$$

1. Because this is the equation set equal to zero in section 5, the determinant of the system must equal zero to determine the values of λ_n.

The ratios of the latent vectors may be computed from any one of the above equations. For example:

$$X_1 = \left\{a_{12}a_{23} - (a_{22} - \lambda_1)\, a_{13}\right\}$$

or:

$$X_1 = \left\{(a_{22} - \lambda_1)(a_{33} - \lambda_1) - a_{23}a_{23}\right\}, \text{ etc.}$$

The latent vectors of the other roots are then solved similarly and the vector matrix X is formed.

$$\begin{bmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{bmatrix} \tag{9}$$

Under the same hypothesis as in the bivariate case, the direction cosines are computed.

$$\begin{array}{c|ccc} & x' & y' & z' \\ \hline x & \cos\alpha_{11} = \dfrac{X_1}{S_1} & \cos\alpha_{12} = \dfrac{X_2}{S_2} & \cos\alpha_{13} = \dfrac{X_3}{S_3} \\ y & \cos\alpha_{21} = \dfrac{Y_1}{S_1} & \cos\alpha_{22} = \dfrac{Y_2}{S_2} & \cos\alpha_{23} = \dfrac{Y_3}{S_3} \\ z & \cos\alpha_{31} = \dfrac{Z_1}{S_1} & \cos\alpha_{32} = \dfrac{Z_2}{S_2} & \cos\alpha_{33} = \dfrac{Z_3}{S_3} \end{array} \tag{10}$$

where:

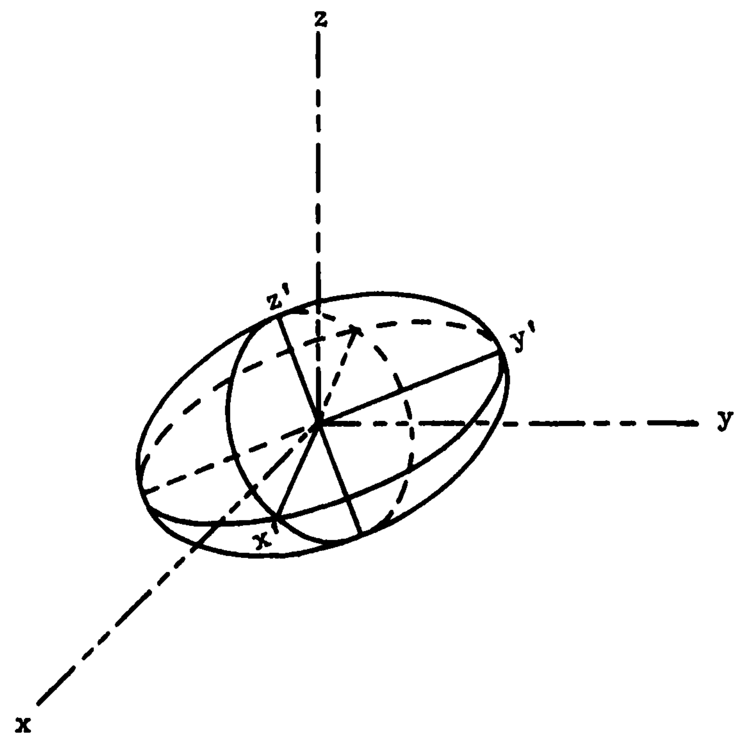
$$S_1 = \sqrt{X_1^2 + Y_1^2 + Z_1^2}$$

$$S_2 = \sqrt{X_2^2 + Y_2^2 + Z_2^2}$$

$$S_3 = \sqrt{X_3^2 + Y_3^2 + Z_3^2}$$

The solution of equation (8) allows for the signs of the vector components, and orientation is similar to that in the bivariate case. That is, the direction cosines, as given in matrix (10), measure the angles between the respective axis.

For example:



As oriented using x', y', z' direction cosines are:

	x'	y'	z'
x	+	-	+
y	+	+	-
z	-	+	+

7.4. Basic Properties of Direction Cosines. Checks resulting from the basic properties of cosines to insure correct values are:

a. $\sum_{n=1}^{2,3} \cos^2\alpha_{n1} = 1; \quad \sum_{n=1}^{2,3} \cos^2\alpha_{n2} = 1; \quad \sum_{n=1}^{2,3} \cos^2\alpha_{n3} = 1$

b. To check the perpendicularity of the principal axes.

Bivariate case:

$$\cos\alpha_{11}\cos\alpha_{21} + \cos\alpha_{12}\cos\alpha_{22} = 0$$

Trivariate case:

$$\cos\alpha_{11}\cos\alpha_{21} + \cos\alpha_{12}\cos\alpha_{22} + \cos\alpha_{13}\cos\alpha_{23} = 0$$
$$\cos\alpha_{11}\cos\alpha_{31} + \cos\alpha_{12}\cos\alpha_{32} + \cos\alpha_{13}\cos\alpha_{33} = 0$$
$$\cos\alpha_{21}\cos\alpha_{31} + \cos\alpha_{22}\cos\alpha_{32} + \cos\alpha_{23}\cos\alpha_{33} = 0$$

c. The determinant of the $[\cos\alpha_{nm}]$ must equal +1

d. Each element in $[\cos\alpha_{nm}]$ must equal its cofactor

e. $X^{-1}AX = \lambda$

By multiplying the inverse of the latent vector matrix by the covariance matrix and the latent vector matrix, a diagonal matrix is obtained which has the latent roots as its diagonal elements.

8. APPLICATION OF THE MOMENT (COVARIANCE) MATRIX TO LEAST SQUARES ADJUSTMENTS INVOLVING CORRELATED OBSERVATIONS

In many geodetic investigations least squares adjustments - a special branch of statistics - are used to obtain consistent estimates of measured variables and to provide estimations of the reliability of the estimates (i.e., standard errors).

This section contains a matrix solution of a general least squares problem, and shows how a system of normal equations is related to the covariance matrix, and how the theory of the preceeding sections are used to reduce a system of correlated observations to one involving derived observations which are uncorrelated.

8.1. Observation Equations. Suppose the following system of observation equations is given:

$$\begin{aligned} a_1x_1 + b_1x_2 + c_1x_3 + \dots m_1x_m + s_1 &= Z_1 \\ a_2x_1 + b_2x_2 + c_2x_3 + \dots m_2x_m + s_2 &= Z_2 \\ a_nx_1 + b_nx_2 + c_nx_3 + \dots m_nx_m + s_n &= Z_n \end{aligned}$$

where:

$a, b, c, \dots, m, s$ are constants known from the theory of the observations, Z is the quantity observed and $x_1, x_2, \dots, x_m$ are unknowns to be found.

If the observations were without error, these equations would be satisfied by the true values $(\xi_1, \xi_2, \dots, \xi_m)$ of $x_1, x_2, \dots, x_m$.

Since all observations are subject to small errors, there is no system of values for the unknowns which satisfies the equations exactly. As previously mentioned, one of the purposes of a least squares adjustment is to find a system of best estimates ($\bar{x}_1$, $\bar{x}_2$, ..., $\bar{x}_m$) for the unknowns.

Let h_1, h_2, h_3, ..., h_n be the values of the left side of the system of observation equations when any assumed set of values (x'_1, x'_2, ..., x'_m) for the unknowns is taken. Then the errors (or more strictly the residuals) of the observations are given by:

$$v_i = h_i - Z_i$$

where:

$$i = 1, 2, \ldots, n$$

Letting:

$$k_i = Z_i - s_i$$

the system of observation equations is written:

$$\begin{array}{c}
a_1x_1 + b_1x_2 + c_1x_3 + \ldots + m_1x_m - k_1 = v_1 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_nx_1 + b_nx_2 + c_nx_3 + \ldots + m_nx_m - k_n = v_n
\end{array}$$

Introducing matrices, let:

$$\underset{n\times m}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

where:

$$\begin{matrix} a_{11} = a_1 \\ a_{21} = a_2 \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{n1} = a_n \end{matrix} \quad , \quad \begin{matrix} a_{12} = b_1 \\ a_{22} = b_2 \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{n2} = b_n \end{matrix} \quad , \text{ etc.}$$

$$\underset{n\times 1}{K} = \begin{bmatrix} k_1 \\ k_2 \\ \cdot \\ \cdot \\ \cdot \\ k_n \end{bmatrix}, \quad \underset{m\times 1}{\bar{X}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_m \end{bmatrix}, \quad \underset{n\times 1}{V} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix}, \quad \underset{m\times 1}{X} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}$$

$$\underset{n\times n}{P} = \begin{bmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & p_n \end{bmatrix} = \text{weight matrix}$$

The residuals of the observations are given by:

$$\underset{n\times 1}{V} = -K + AX$$

If X has the value $\bar{X}$,

$$(1) \quad V = -K + A\bar{X}$$

8.2. <u>The Theory of Least Squares</u>. The theory of least squares requires that the sum of the squares of the weighted residuals be a minimum, that is $[pv^2]$ or $[pvv]$ equals a minimum. A necessary condition for the fulfilment of this requirement is that the m partial derivatives of $[pvv]$ with respect to $\bar{x}_1, \ldots, \bar{x}_m$ be zero:

$$(2) \quad \frac{\partial}{\partial \bar{x}_j} [pvv] = 2 \left(p_1 v_1 \frac{\partial v_1}{\partial \bar{x}_j} + \ldots p_n v_n \frac{\partial v_n}{\partial \bar{x}_j} \right) = 0$$

where:

$$j = 1, 2, \ldots, m$$

From (1):

$$\frac{\partial v_i}{\partial x_j} = [a_{ij}] = A$$

where:

$$i = 1, 2, \ldots, n$$

$$j = 1, 2, \ldots, m$$

An alternate way of writing (2) is:

$$A^T \cdot P \cdot V = 0$$

Inserting (1) into this equation yields:

$$A^T \cdot P \cdot (- K + A\bar{X}) = 0$$

or:

$$A^T \cdot P \cdot A\bar{X} = A^T \cdot P \cdot K$$

8.3. Solution. The matrix product $A^T \cdot P \cdot A$ is equivalent to B, the symmetric coefficient matrix of the normal equations for the system of observation equations. Thus the preceeding expression may be written:

$$B\bar{X} = A^T \cdot P \cdot K$$

Letting $A^T \cdot P \cdot K$ equal C, the solution for the unknowns in matrix notation is:

$$\bar{X} = B^{-1}C$$

8.4. Verification. [pvv] can be obtained from the matrix product:

$$[pvv] = V^T \cdot P \cdot V$$

To verify that V actually minimizes [pvv], consider another set of corrections V' and values X'.

In this case:

$$V' = -K + AX'$$

Subtracting V from V':

$$V' - V = A\,(X' - \bar{X})$$

Now:

$$V' = (V' - V) + V$$

and:

$$[pv'v'] = V'^{T} \cdot P \cdot V'$$

Substituting $V' = (V' - V) + V$:

$$[pv'v'] = \Big[(V' - V) + V\Big]^{T} \cdot P \cdot \Big[(V' - V) + V\Big]$$

Applying the distribution laws of matrix multiplication:

$$[pv'v'] = (V' - V)^{T} \cdot P \cdot (V' - V) + (V' - V)^{T} \cdot P \cdot V$$

$$+ V^{T} \cdot P \cdot (V' - V) + V^{T} \cdot P \cdot V$$

rearranging terms:

$$[pv'v'] = V^{T} \cdot P \cdot V + (V' - V)^{T} \cdot P \cdot (V' - V) + (V' - V)^{T} \cdot P \cdot V$$

$$+ V^{T} \cdot P \cdot (V' - V)$$

Substituting $(V' - V) = A\,(X' - \bar{X})$ into the third and fourth terms:

$$[pv'v'] = V^T \cdot P \cdot V + (V' - V)^T \cdot P \cdot (V' - V) + \left\{A\,(X' - \bar{X})\right\}^T \cdot P \cdot V$$
$$+ V^T \cdot P \cdot \left\{A \cdot (X' - \bar{X})\right\}$$

$$= V^T \cdot P \cdot V + (V' - V)^T \cdot P \cdot (V' - V) + (X' - \bar{X})^T \cdot A^T \cdot P \cdot V$$
$$+ V^T \cdot P \cdot A \cdot (X' - \bar{X})$$

but:

$$A^T \cdot P \cdot V = 0$$

thus

$$V^T \cdot P^T \cdot A = 0^T = 0$$

and

$$[pv'v'] = V^T \cdot P \cdot V + (V' - V)^T \cdot P \cdot (V' - V) + 0 + 0$$

$$= [pvv] + \left[\,p\,(v' - v)(v' - v)\right]$$

$$= [pvv] + \left[\,p\,(v' - v)^2\right]$$

The factor $\left[\,p\,(v' - v)^2\right]$ is positive or equal to zero. Thus it is seen that [pv'v'] is a minimum when $\left[\,p\,(v' - v)^2\right]$ is zero or when $V' = V$.

8.5. <u>The Moment Matrix.</u> Now assume that the true values ξ_1, ξ_2, ..., ξ_m and true errors ϵ_1, ϵ_2, ..., ϵ_m were known.

Introducing:

$$E = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdot \\ \cdot \\ \cdot \\ \xi_m \end{bmatrix} \quad \text{and} \quad \epsilon' = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_m \end{bmatrix}$$

where:

$$\epsilon_i = \bar{X}_i - \xi_i$$

$$[p\epsilon\epsilon] = \text{the sum of the squares of the weighted true errors}$$

$$= [pvv] + p\,(\epsilon' - V)^2$$

$$= [pvv] + (\epsilon' - V)^T \cdot P \cdot (\epsilon' - V)$$

but:

$$\epsilon' - V = A\,(E - \bar{X})$$

$$\therefore [p\epsilon\epsilon] = [pvv] + \left\{A \cdot (E - \bar{X})\right\}^T \cdot P \cdot \left\{A \cdot (E - \bar{X})\right\}$$

$$= [pvv] + (E - \bar{X})^T \cdot A^T \cdot P \cdot A\,(E - \bar{X})$$

but:

$$A^T \cdot P \cdot A = B$$

and:

$$[p\epsilon\epsilon] = [pvv] + (E - \bar{X})^T \cdot B \cdot (E - \bar{X})$$

$$= q^2 + (E - \bar{X})^T \cdot B \cdot (E - \bar{X})$$

where:

$$q = \sqrt{[pvv]}$$

From statistical theory it is known that the normal error distribution of $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m$ and q can be written in the form:

$$f\,(\bar{x}_1, \ldots, \bar{x}_m, q) = K \exp\left\{- \frac{1}{2} Q\right\}$$

where:

(3) Q (quadratic form of the error distribution)

$$Q = \frac{(\bar{X} - E)^T \cdot B \cdot (\bar{X} - E)}{\sigma^2} \qquad \text{(ref. 1)}$$

σ^2 = the variance of a single observation after adjustment.

and:

K is a constant corresponding to Q.

It was shown in section 4.2. that Q of a normal bivariate error distribution is:

$$Q = [X\ Y]^T \cdot M^{-1} \cdot [X\ Y]$$

where:

M^{-1} = the inverse of the moment matrix.

Comparing this expression to (3) indicates that:

$$M^{-1} = \frac{B}{\sigma^2}$$

Multiplying from the left by M:

$$MM^{-1} = \frac{MB}{\sigma^2}$$

$$I = \frac{MB}{\sigma^2}$$

Multiplying from the right by B^{-1}:

$$IB^{-1} = \frac{MBB^{-1}}{\sigma^2}$$

$$B^{-1} = \frac{M}{\sigma^2}$$

or:

$$M = \sigma^2 B^{-1}$$

The latter equation states that the moment matrix (or covariance matrix) of the best estimates of the unknowns is equivalent to the product of the variance of a single observation and the inverse of the coefficient matrix of the normal equations.

M in the multivariate case is:

$$\underset{m \times m}{M} = \begin{bmatrix} \sigma^2_{x_1} & \rho_{x_1 x_2}\sigma_{x_1}\sigma_{x_2} & \cdots & \rho_{x_1 x_m}\sigma_{x_1}\sigma_{x_m} \\ & & \cdots & \\ \rho_{x_1 x_m}\sigma_{x_1}\sigma_{x_m} & \rho_{x_2 x_m}\sigma_{x_2}\sigma_{x_m} & \cdots & \sigma^2_{x_m} \end{bmatrix}$$

where the diagonal elements are variances, the non-diagonal elements covariances, and the ρ's correlation coefficients.

8.6. Transformation of Errors. In the case of independent (uncorrelated) errors, the ρ's are zero and M is a diagonal matrix or equivalently B^{-1} is a diagonal matrix; in the case of dependent (correlated) errors the ρ's are not zero and B^{-1} is not a diagonal matrix. Thus it may be concluded that: if B^{-1}, for correlated observations, is transformed into a diagonal matrix, then uncorrelated errors are obtained.

To effect such a transformation the process of diagonalization is used (section 5 for the bivariate case). For the multivariate case, the determinant of B^{-1}, with a factor λ subtracted from each diagonal term, is expanded and set equal to zero. These operations yield an m^{th} degree characteristic equation having m real roots. The roots when multiplied by σ^2 yield the variances of the unknowns in an uncorrelated sense. The standard errors are found by extracting the square roots of the variances.

APPENDIX A

EXPANSION OF THE NORMAL BIVARIATE MOMENT GENERATING FUNCTION

(Reference No. 12)

In section 4.3., it was asserted that:

$$E\left\{\exp\left[\theta x + \theta y\right]\right\} = \frac{\sqrt{A}}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2}\left(a_{11}x^2 + 2a_{12}xy + a_{22}y^2\right) + \theta_x x + \theta_y y\right\}dxdy \qquad (1)$$

$$= \frac{\sqrt{A}}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2}a_{11}\left(x + \frac{a_{12}}{a_{11}}y - \frac{\theta_x}{a_{11}}\right)^2 - \frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)\left(y + \frac{a_{12}\theta_x - a_{11}\theta_y}{|A|}\right)^2 + \frac{1}{2}R\right\}dxdy \qquad (2)$$

where:

$$R = \frac{a_{22}\theta_x^2 + a_{11}\theta_y^2 - 2a_{12}\theta_x\theta_y}{|A|}$$

Let:

$$\alpha = 2x \left[\text{the exponential in } (1)\right]$$

$$= -a_{11}x^2 - 2a_{12}xy - a_{22}y^2 + 2\theta_x x + 2\theta_y y$$

and:

$$\beta_1 = -a_{11}\left(x + \frac{a_{12}y}{a_{11}} - \frac{\theta_x}{a_{11}}\right)^2$$

$$\beta_2 = -\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)\left(y + \frac{a_{12}\theta_x}{|A|} - \frac{a_{11}\theta_y}{|A|}\right)^2$$

To prove that (1) equals (2) it will suffice to show that:

$$\alpha = \beta_1 + \beta_2 + R$$

or equivalently that:

$$R = \alpha - \beta_1 - \beta_2$$

$$\beta_1 = -a_{11}\left(x + \frac{a_{12}}{a_{11}}y - \frac{\theta_x}{a_{11}}\right)^2$$

$$= -a_{11}\left(x^2 + \frac{a_{12}}{a_{11}}xy - \frac{\theta_x}{a_{11}}x + \frac{a_{12}}{a_{11}}xy + \frac{a_{12}^2}{a_{11}^2}y^2 - \frac{a_{12}\theta_x y}{a_{11}^2} - \frac{\theta_x x}{a_{11}} - \frac{a_{12}\theta_x y}{a_{11}^2} + \frac{\theta_x^2}{a_{11}^2}\right)$$

$$\beta_1 = -a_{11}x^2 - 2a_{12}xy + 2\theta_x x + \frac{2a_{12}\theta_x y}{a_{11}} - \frac{a_{12}^2 y^2}{a_{11}} - \frac{\theta_x^2}{a_{11}}$$

$$\beta_2 = -\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)\left(y + \frac{a_{12}\theta_x}{|A|} - \frac{a_{11}\theta_y}{|A|}\right)^2$$

$$= -\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)\left(y^2 + \frac{2a_{12}\theta_x y}{|A|} - \frac{2a_{11}\theta_y y}{|A|} - \frac{2a_{11}a_{12}\theta_x\theta_y}{|A|^2} + \frac{a_{12}^2\theta_x^2}{|A|^2} + \frac{a_{11}^2\theta_y^2}{|A|^2}\right)$$

Expanding:

$$\beta_2 = -a_{22}y^2 + \frac{a_{12}^2y^2}{a_{11}} - \frac{2a_{12}a_{22}\theta_x y}{|A|} + \frac{2a_{12}^3\theta_x y}{|A|\,a_{11}} + \frac{2a_{11}a_{22}\theta_y y}{|A|} - \frac{2a_{11}a_{12}^2\theta_y y}{|A|\,a_{11}}$$

$$+ \frac{2a_{11}a_{12}a_{22}\theta_x\theta_y}{|A|^2} - \frac{2a_{11}a_{12}^3\theta_x\theta_y}{|A|^2a_{11}} - \frac{a_{12}^2a_{22}\theta_x^2}{|A|^2} + \frac{a_{12}^4\theta_x^2}{|A|^2a_{11}} - \frac{a_{11}^2a_{22}\theta_y^2}{|A|^2}$$

$$+ \frac{a_{11}^2a_{12}^2\theta_y^2}{|A|^2a_{11}}$$

Collecting terms:

$$-R = -\alpha + \beta_1 + \beta_2$$

$$-R = \Bigg(a_{11}x^2 + 2a_{12}xy + a_{22}y^2 - 2\theta_x x - 2\theta_y y - a_{11}x^2 - 2a_{12}xy + \frac{2a_{12}\theta_x y}{a_{11}}$$

$$+ 2\theta_x x - \frac{a_{12}y^2}{a_{11}} - \frac{\theta_x^2}{a_{11}} - a_{22}y^2 + \frac{a_{12}^2y^2}{a_{11}} - \frac{2a_{12}a_{22}\theta_x y}{|A|} + \frac{2a_{12}^3\theta_x y}{|A|\,a_{11}}$$

$$+ \frac{2a_{11}a_{22}\theta_y y}{|A|} - \frac{2a_{11}a_{12}^2\theta_y y}{|A|\,a_{11}} + \frac{2a_{11}a_{12}a_{22}\theta_x\theta_y}{|A|^2} - \frac{2a_{11}a_{12}^3\theta_x\theta_y}{|A|^2a_{11}}$$

$$- \frac{a_{12}^2a_{22}\theta_x^2}{|A|^2} + \frac{a_{12}^4\theta_x^2}{|A|^2a_{11}} - \frac{a_{11}^2a_{22}\theta_y^2}{|A|^2} + \frac{a_{11}^2a_{12}^2\theta_y^2}{|A|^2a_{11}} \Bigg)$$

Combining terms (small numbers circled above expressions indicates expressions combined)

$$\textcircled{1} \qquad a_{11}x^2 - a_{11}x^2 = 0$$

$$\textcircled{2} \qquad 2a_{12}xy - 2a_{12}xy = 0$$

$$\textcircled{3} \qquad + a_{22}y^2 - a_{22}y^2 = 0$$

$$\textcircled{4} \qquad + 2\theta_x x - 2\theta_x x = 0$$

$$\textcircled{5} \qquad + \frac{2a_{12}\theta_x y}{a_{11}} - \frac{2a_{12}a_{22}\theta_x y}{|A|} + \frac{2a_{12}^3\theta_x y}{|A|\,a_{11}}$$

$$= \frac{2|A|\,a_{12}\theta_x y - 2a_{12}a_{11}a_{22}\theta_x y + 2a_{12}^3\theta_x y}{|A|\,a_{11}}$$

$$= \frac{2|A|\,a_{12}\theta_x y - 2a_{12}\theta_x y\left(a_{11}a_{22} - a_{12}^2\right)}{|A|\,a_{11}}$$

Since:

$$a_{11}a_{22} - a_{12}^2 = |A|$$

$$= \frac{2a_{12}\theta_x y - 2a_{12}\theta_x y}{a_{11}} = 0$$

$$(6) \quad - \frac{a_{12}^2 y^2}{a_{11}} + \frac{a_{12}^2 y^2}{a_{11}} = 0$$

$$(7) \quad + \frac{2a_{11}a_{22}\theta_y y}{|A|} - \frac{2a_{11}a_{12}^2\theta_y y}{|A|\,a_{11}}$$

Since:

$$a_{11}a_{22} = |A| + a_{12}^2$$

$$= \frac{\left(2|A| + 2a_{12}^2 - 2a_{12}^2\right)\theta_y y}{|A|} = 2\theta_y y$$

$$(7) + (8) \qquad - 2\theta_y y + 2\theta_y y = 0$$

$$(9) \quad + \frac{2a_{11}a_{12}a_{22}\theta_x\theta_y}{|A|^2} - \frac{2a_{11}a_{12}^3\theta_x\theta_y}{|A|^2 a_{11}} = \frac{\left(2a_{11}a_{12}a_{22} - 2a_{12}^3\right)\theta_x\theta_y}{|A|^2}$$

$$= \frac{2a_{12}\left(a_{11}a_{22} - a_{12}^2\right)\theta_x\theta_y}{|A|^2} = \frac{2a_{12}\theta_x\theta_y}{|A|}$$

$$(10) \quad - \frac{a_{12}^2 a_{22}\theta_x^2}{|A|^2} + \frac{a_{12}^4\theta_x^2}{|A|^2 a_{11}} - \frac{\theta_x^2}{a_{11}} = - \frac{|A|^2\theta_x^2 - a_{11}a_{22}a_{12}^2\theta_x^2 + a_{12}^4\theta_x^2}{|A|^2 a_{11}}$$

$$= - \frac{|A|^2\theta_x^2 - a_{12}^2\theta_x^2\left(a_{11}a_{22} - a_{12}^2\right)}{|A|^2 a_{11}} = - \frac{|A|^2\theta_x^2 - |A|\,a_{12}^2\theta_x^2}{|A|^2 a_{11}}$$

$$= -\frac{\theta_x^2\left(|A| + a_{12}^2\right)}{|A|\, a_{11}} = -\frac{\theta_x^2\left(a_{11}a_{22} - a_{12}^2 + a_{12}^2\right)}{|A|\, a_{11}} = -\frac{a_{22}\theta_x^2}{|A|}$$

$$⑪ \quad -\frac{a_{11}^2 a_{22}\theta_y^2}{|A|^2} + \frac{a_{11}^2 a_{12}^2\theta_y^2}{|A|^2 a_{11}} = -\frac{a_{11}^3 a_{22}\theta_y^2 + a_{11}^2 a_{12}^2\theta_y^2}{a_{11}|A|^2}$$

$$= -\frac{\theta_y^2 a_{11}\left(a_{11}a_{22} - a_{12}^2\right)}{|A|^2} = -\frac{a_{11}\theta_y^2}{|A|}$$

Therefore:

$$-R = ⑨ + ⑩ + ⑪$$

$$-R = \frac{2a_{12}\theta_x\theta_y}{|A|} - \frac{a_{22}\theta_x^2}{|A|} - \frac{a_{11}\theta_y^2}{|A|}$$

or:

$$R = \frac{a_{22}\theta_x^2 - 2a_{12}\theta_x\theta_y + a_{11}\theta_y^2}{|A|}$$

Since

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \frac{a_{22}}{|A|} & -\frac{a_{12}}{|A|} \\ -\frac{a_{12}}{|A|} & \frac{a_{11}}{|A|} \end{bmatrix} = 1 \equiv \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} a^{11} & a^{12} \\ a^{12} & a^{22} \end{bmatrix}$$

Comparison yields

$$\frac{a_{22}}{|A|} = a^{11}$$

$$\frac{a_{12}}{|A|} = -a^{12}$$

etc.

Thus: R may also be written as:

$$R = a^{11}\theta_x^2 + a^{22}\theta_y^2 + 2a^{12}\theta_x\theta_y$$

APPENDIX B

DEVELOPMENT OF THE CONSTANT FOR A TRIVARIATE ERROR DISTRIBUTION

(Reference No. 12)

This development can be carried out by introducing a change in the designation of the coordinate axes and following the procedure used in the bivariate case.

From section 6.

$$f(x,y,z) = K_3 \exp\left\{-\frac{1}{2} Q_3\right\}$$

Taking the integral over the entire universe:

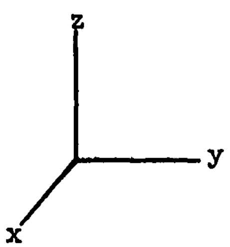
$$F(x,y,z) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} K_3 \exp\left\{-\frac{1}{2} Q_3\right\} dxdydz = 1$$

where:

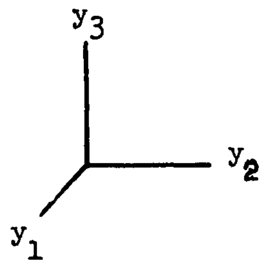
$$Q_3 = a_{11}x^2 + 2a_{12}xy + 2a_{13}xz + a_{22}y^2 + 2a_{23}yz + a_{33}z^2$$

$$\therefore \frac{1}{K_3} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2} Q_3\right\} dxdydz$$

Making the coordinate notation change:



to



Q_3 can now be written in simplified form by:

$$Q_3 = \sum_{i,j=1}^{3} a_{ij} y_i y_j$$

where:

$$y_1 = x$$

$$y_2 = y$$

$$y_3 = z$$

Q_3 is now written:

$$a_{11}\left(y_1 + \frac{\sum_{j=2}^{3} a_{1j} y_j}{a_{11}}\right)^2 + \sum_{i,j=2}^{3}\left(a_{ij} - \frac{a_{1i} a_{1j}}{a_{11}}\right) y_i y_j$$

Letting:

$$Z_1 = y_1 + \frac{\sum_{j=2}^{3} a_{1j} y_j}{a_{11}}$$

$$Z_2 = y_2 + \frac{a'_{23} y_3}{a'_{22}}$$

$$Z_3 = y_3$$

$$a'_{ij} = a_{ij} - \frac{a_{1j} a_{i1}}{a_{11}} \text{ (ref. 11)}$$

$$a_{ij}^{2} = a_{ij}^{1} - \frac{a_{2j}^{1} a_{i2}^{1}}{a_{22}^{1}}$$

$$a_{kk}^{k-1} = a_{kk}^{k-2} - \frac{a_{k-1,k}^{k-2} a_{k,k-1}^{k-2}}{a_{k-1,k-1}^{k-2}}$$

Using these substitutions the integral becomes:

$$\frac{1}{K_3} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2} a_{11}Z_1^2 - \frac{1}{2} a_{22}^{1}Z_2^2 - \frac{1}{2} a_{33}^{2}Z_3^2\right\} dZ_1 dZ_2 dZ_3$$

Next it will be shown that:

$$Q_3 = a_{11}Z_1^2 - \frac{1}{2} a_{22}^{1}Z_2^2 - \frac{1}{2} a_{33}^{2}Z_3^2$$

Once this has been proven, K_3 can be found by integration.

Proof: In a symmetric quadratic form:

$$a_{ij} = a_{ji}$$

$$\therefore a_{11} = a_{11}$$

$$a_{12} = a_{21}$$

$$\vdots \qquad \vdots$$

$$\underset{\substack{i=2\\j=2}}{a_{22}^{1}} = a_{22} - \frac{a_{12}^2}{a_{11}}$$

$$a''_{33} \underset{\substack{i=3\\ j=3}}{} = a'_{33} - \frac{a'_{23}a'_{23}}{a'_{22}}$$

$$a'_{33} \underset{\substack{i=3\\ j=3}}{} = a_{33} - \frac{a_{13}^2}{a_{11}}$$

$$a'_{23} \underset{\substack{i=2\\ j=3}}{} = a_{23} - \frac{a_{13}a_{21}}{a_{11}}$$

$$Z_1 = y_1 + \frac{a_{12}}{a_{11}} y_2 + \frac{a_{13}}{a_{11}} y_3$$

$$Z_2 = y_2 + \frac{a'_{23}}{a'_{22}} y_3$$

$$Z_3 = y_3$$

Using the above substitutions:

$$\frac{1}{2} a_{11}Z_1^2 = \frac{1}{2}\left(a_{11}y_1^2 + 2a_{12}y_1y_2 + 2a_{13}y_1y_3 + \overbrace{\frac{a_{12}^2y_2^2}{a_{11}}}^{1} + \overbrace{\frac{2a_{12}a_{13}y_2y_3}{a_{11}}}^{2} + \overbrace{\frac{a_{13}^2y_3^2}{a_{11}}}^{3}\right)$$

and:

$$Z_2 = y_2 + \frac{a'_{23}}{a'_{22}} y_3 = y_2 + \frac{1}{a'_{22}}\left(a_{23} - \frac{a_{13}a_{21}}{a_{11}}\right) y_3$$

$$\frac{1}{2} a'_{22} Z_2^2 = \frac{1}{2}\left\{\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right) y_2^2 + 2\left(a_{11}a_{23} - \frac{a_{12}a_{13}}{a_{11}}\right) y_2 y_3 + \frac{1}{a'_{22}}\left(\frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}}\right)^2 y_3^2\right\}$$

$$= \frac{1}{2}\left\{a_{22} y_2^2 - \overbrace{\frac{a_{12}^2}{a_{11}}}^{①} y_2^2 + 2 y_2 y_3 \frac{\overbrace{\left(a_{11}a_{23} - a_{12}a_{13}\right)}^{②}}{a_{11}} + \overbrace{\frac{1}{a'_{22}}\left(\frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}}\right)^2 y_3^2}^{④}\right\}$$

Also:

$$\frac{1}{2} a'_{33} Z_3^2 = \left\{ a_{33} y_3^2 - \overbrace{\frac{a_{13}^2 y_3^2}{a_{11}}}^{③} - \overbrace{\frac{1}{a'_{22}}\left(\frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}}\right)^2 y_3^2}^{④}\right\}$$

Adding $\frac{1}{2}\left\{a_{11} Z_1^2 + a'_{22} Z_2^2 + a'_{33} Z_3^2\right\}$ and combining like terms yields:

$$\frac{1}{2}\left\{a_{11} y_1^2 + 2a_{12} y_1 y_2 + 2a_{13} y_1 y_3 + a_{22} y_2^2 + 2a_{23} y_2 y_3 + a_{33} y_3^2\right\}$$

The expression within the bracket is Q. QED

Now:

$$\int_{-\infty}^{+\infty} \exp\left\{- cx^2\right\} dx = \sqrt{\frac{\pi}{c}}$$

applying this integral to the integral for $\frac{1}{K_3}$ yields:

$$\frac{1}{K_3} = \int_{-\infty}^{+\infty} \exp\left\{- \frac{1}{2} a_{11} Z_1^2\right\} dZ_1 \cdot \int_{-\infty}^{+\infty} \exp\left\{- \frac{1}{2} a_{22}' Z_2^2\right\} dZ_2 \cdot \int_{-\infty}^{+\infty} \exp\left\{- \frac{1}{2} a_{33}'' Z_3^2\right\} dZ_3$$

$$\frac{1}{K_3} = \sqrt{\frac{\pi}{\frac{1}{2} a_{11}} \cdot \frac{\pi}{\frac{1}{2} a_{22}'} \cdot \frac{\pi}{\frac{1}{2} a_{33}''}} = \sqrt{\frac{(2\pi)^3}{a_{11} \cdot a_{22}' \cdot a_{33}''}}$$

$\sqrt{a_{11} \cdot a_{22}' \cdot a_{33}''}$ can be written in a more convenient form.

Since:

$$a_{11} = a_{11}$$

$$a_{22}' = a_{22} - \frac{\left(a_{12}'\right)^2}{a_{11}}$$

$$a_{33}'' = a_{33}' - \frac{a_{23}' a_{23}'}{a_{22}'}$$

The product of the radicand is:

$$\left(a_{11}\right) \left(a_{22}'\right) \left| a_{33} - \frac{\left(a_{23}'\right)^2}{a_{22}'} \right| = \left| a_{11} \cdot a_{22}' \cdot a_{33}' - a_{11} a_{22}' \frac{\left(a_{23}'\right)^2}{a_{22}'} \right.$$

$$= \left\{ a_{11} \cdot a_{22}^{1} \cdot a_{33}^{1} - a_{11} \left(a_{23}^{1} \right)^{2} \right\} = a_{11} \left\{ a_{22}^{1} \cdot a_{33}^{1} - \left(a_{23}^{1} \right)^{2} \right.$$

$$= a_{11} \left| \left(a_{22} - \frac{a_{12}^{2}}{a_{11}} \right) \left(a_{33} - \frac{a_{13}^{2}}{a_{11}} \right) - a_{23}^{2} + \frac{2a_{23}a_{13}a_{12}}{a_{11}} - \frac{a_{13}^{2}a_{12}^{2}}{a_{11}^{2}} \right|$$

Expanding:

$$= a_{11}a_{22}a_{33} - a_{22}a_{13}^{2} - a_{33}a_{12}^{2} - a_{11}a_{23}^{2} + 2a_{23}a_{21}a_{31}$$

This expression is merely the expansion of the 3×3 symmetric determinant $|A|$:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$$

Therefore:

$$\frac{1}{K_3} = \sqrt{\frac{(2\pi)^3}{|A|}}$$

or:

$$K_3 = \sqrt{\frac{|A|}{(2\pi)^3}}$$

APPENDIX C

PROOF OF ASSUMPTION THAT $\mu_{xy} = \rho\sigma_x\sigma_y$ IN NORMAL BIVARIATE ERROR DISTRIBUTION

In section 4.3. the general normal bivariate error distribution was derived using the assumption that ($\mu_{xy} = \rho\sigma_x\sigma_y$). Determination of the mixed product moment, (μ_{xy}) should equal $\rho\sigma_x\sigma_y$ if the assumption is correct.

From section 4.3.

$$f(x,y)dxdy = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)}\left(\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2} - \frac{2\rho(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} \right)\right\} dxdy$$

To prove: $\mu_{xy} = \rho\sigma_x\sigma_y$

$$\mu_{xy} = E\left\{(x-\bar{x})(y-\bar{y})\right\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\bar{x})(y-\bar{y})\ f(x,y)dxdy$$

Let:

$$\frac{(x-\bar{x})}{\sigma_x} = Z_1 \quad \text{and} \quad \frac{(y-\bar{y})}{\sigma_y} = Z_2$$

Then:

$$x - \bar{x} = Z_1\sigma_x \qquad y - \bar{y} = Z_2\sigma_y$$

$$x = Z_1\sigma_x + \bar{x} \qquad y = Z_2\sigma_y + \bar{y}$$

$$dx = \sigma_x dZ_1 \qquad dy = \sigma_y dZ_2$$

Then substituting in:

$$\mu_{xy} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{Z_1\sigma_x Z_2\sigma_y}{2\pi\sigma_x\sigma_y\sqrt{1 - \rho^2}} \exp\left\{- \frac{1}{2\,(1 - \rho^2)}\left(Z_1^2 - 2\rho Z_1 Z_2 + Z_2^2\right)\right\} \sigma_x dZ_1 \sigma_y dZ_2$$

For a given distribution the quantities σ_x, σ_y, ρ and 2π are constants:

$$\mu_{xy} = \frac{\sigma_x\sigma_y}{2\pi\sqrt{(1 - \rho^2)}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Z_1 Z_2 \exp\left\{- \frac{1}{2\,(1 - \rho^2)}\left(Z_1^2 - 2\rho Z_1 Z_2 + Z_2^2\right)\right\} dZ_1 dZ_2$$

Introducing two new variables:

$$Z_1 = \frac{v - w}{\sqrt{2}} \quad ; \quad Z_2 = \frac{v + w}{\sqrt{2}} \quad ; \quad Z_1 Z_2 = \frac{v^2 - w^2}{2}$$

Then:

$$Z_1^2 - 2\rho Z_1 Z_2 + Z_2^2 = \frac{v^2}{2} - \frac{2vw}{2} + \frac{w^2}{2} - 2\rho\left(\frac{v^2}{2} - \frac{w^2}{2}\right) + \frac{v^2}{2} + \frac{2vw}{2} + \frac{w^2}{2}$$

$$= v^2\,(1 - \rho) + w^2\,(1 + \rho)$$

Since the Jacobian of this transformation is unity, $dZ_1 dZ_2$ equals $dvdw$.

Now $(1 - \rho)(1 + \rho) = 1 - \rho^2$

Thus:

$$\mu_{xy} = \frac{\sigma_x\sigma_y}{2\pi\sqrt{1 - \rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{v^2 - w^2}{2} \exp\left\{- \frac{1}{2}\left(\frac{v^2\,(1 - \rho)}{(1 - \rho)(1 + \rho)} + \frac{w^2\,(1 + \rho)}{(1 - \rho)(1 + \rho)}\right)\right\} dvdw$$

$$\mu_{xy} = \frac{\sigma_x \sigma_y}{2} \left| \frac{1}{2\pi} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} v^2 \exp\left\{- \frac{1}{2}\left(\frac{v^2}{1+\rho} + \frac{w^2}{1-\rho}\right)\right\} \frac{dv}{\sqrt{1+\rho}} \frac{dw}{\sqrt{1-\rho}} \right.$$

$$\left. - \frac{1}{2\pi} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} w^2 \exp\left\{- \frac{1}{2}\left(\frac{v^2}{1+\rho} + \frac{w^2}{1-\rho}\right)\right\} \frac{dv}{\sqrt{1+\rho}} \frac{dw}{\sqrt{1-\rho}} \right|$$

The first integral can be written:

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} v^2 \exp\left\{- \frac{1}{2}\left(\frac{v^2}{1+\rho} + \frac{w^2}{1-\rho}\right)\right\} \frac{dv}{\sqrt{1+\rho}} \frac{dw}{\sqrt{1-\rho}}$$

or:

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left\{- \frac{1}{2}\left(\frac{w^2}{1-\rho}\right)\right\} \frac{dw}{\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} v^2 \exp\left\{- \frac{1}{2}\left(\frac{v^2}{1+\rho}\right)\right\} dv$$

Now:

$$\int_{-\infty}^{+\infty} v^2 \exp\left\{- \frac{1}{2}\left(\frac{v^2}{1+\rho}\right)\right\} dv \text{ is of the form } \int_{-\infty}^{+\infty} x^2 \exp\left\{- cx^2\right\} dx$$

which is

$$2\left(\frac{1}{2^2 c} \sqrt{\frac{\pi}{c}}\right)$$

where: $c = \dfrac{1}{2\,(1+\rho)}$

Thus the first integral becomes:

$$\int_{-\infty}^{+\infty} \exp\left\{- \frac{1}{2}\left(\frac{w^2}{1-\rho}\right)\right\} \frac{dw}{\sqrt{1+\rho}\sqrt{1-\rho}} \cdot \frac{1}{2\pi} \sqrt{\pi}\ \sqrt{2}\sqrt{(1+\rho)^3}$$

or:

$$\frac{1+\rho}{\sqrt{1-\rho}}\frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{2}}\,2\int_{-\infty}^{+\infty}\exp\left\{-\frac{w^2}{2\,(1-\rho)}\right\}dw$$

The latter expression is of the form:

$$\int_{-\infty}^{+\infty}\exp\left\{-c^2x^2\right\}dx = \frac{\sqrt{\pi}}{2c}$$

with:

$$c^2 = \frac{1}{2\,(1-\rho)}$$

and:

$$c = \frac{1}{\sqrt{2}\,\sqrt{1-\rho}}$$

Therefore, integration of the first integral yields:

$$\frac{1+\rho}{\sqrt{1-\rho}}\frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{2}}\,2\cdot\frac{\sqrt{\pi}}{2}\sqrt{2}\,\sqrt{1-\rho} = 1+\rho$$

Using similar methods on the second integral yield $-(1 - \rho)$ as the result.

Therefore:

$$\mu_{xy} = \frac{\sigma_x\sigma_y}{2}\ (1+\rho) - (1-\rho)$$

or:

$$\mu_{xy} = \rho\sigma_x\sigma_y$$

REFERENCES

1. Aeronautical Chart and Information Center Technical Report No. 96, "Principles of Error Theory and Cartographic Applications", February 1962.

2. Arley, Niels, and Buch, K. Rander, Introduction to the Theory of Probability and Statistics, New York: John Wiley and Son Inc., 1950.

3. Brown, Duane, "A Matrix Treatment of the General Problem of Least Squares Considering Correlated Observations", Ballistic Research Laboratories Report No. 937, May 1955.

4. Cramer, H., Mathematical Methods of Statistics, Stockholm and Princeton: Princeton University Press, 1946.

5. Faddeeva, V.N., (Translated by Curtis D. Benster), Computational Methods of Linear Algebra, New York: Dover Publication Inc., 1959.

6. Eisenhart, Luther Pfahler, Coordinate Geometry, Boston, New York: Ginn and Co., 1939.

7. Hirvonen, R.A., Adjustment Computation Lecture Notes, O.S.U. 1960.

8. Hoel, Paul G., Introduction to Mathematical Statistics, London: John Wiley and Son, Inc.

9. Hotelling, Harold, "Analysis of a Complex of Statistical Variables Into Principal Components", The Journal of Educational Psychology Vol. 24, 1933, pp. 417-441.

10. Michael, Aristotle D., Matrix and Tensor Calculus with Applications to Mechanics, Elasticity and Aeronautics, John Wiley and Son Inc., 1947.

11. Perlis, Sam, Theory of Matrices, Cambridge Massachusetts: Addison Wesley Publishing Company, Inc.

12. Wilks, S.S., Mathematical Statistics, New York: John Wiley and Son, 1962.